

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 04:10:34 ; Search time 160 Seconds
(without alignments)
6381.479 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 cgcgcgctntaagcattt.....tccacagctgctgcgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	566.2	90.7	2052	3	US-08-630-916A-45
C 2	566.2	90.7	2052	3	Sequence 45, Appli
C 3	246.6	39.5	2559	2	Sequence 3, Appli
C 4	246.6	39.5	2559	3	Sequence 2, Appli
C 5	246.6	39.5	2678	4	Sequence 234, App
C 6	246.6	39.5	5372	3	Sequence 1, Appli
C 7	246.2	39.5	5359	2	Sequence 1, Appli
C 8	203.6	32.6	3475	3	Sequence 10, Appl
C 9	203.6	32.6	3476	3	Sequence 47, Appl
C 10	111.8	17.9	2351	4	Sequence 100, App
C 11	54	8.7	2790	3	Sequence 2, Appli
C 12	53.8	8.6	2264	4	Sequence 1035, Ap
C 13	53.4	8.6	2848	3	Sequence 3, Appli
C 14	53.4	8.6	2848	4	Sequence 3, Appli
C 15	52.2	8.4	267	4	Sequence 4083, Ap
C 16	50	8.0	3226	3	Sequence 5, Appli
C 17	50	8.0	3226	4	Sequence 239, App
C 18	48.4	7.8	200	4	Sequence 714, App
C 19	48.4	7.8	4899	4	Sequence 4084, Ap
C 20	47.6	7.6	1395	4	Sequence 49, Appl
C 21	45.2	7.2	673	3	Sequence 3075, Ap
C 22	45.2	7.2	892	4	Sequence 2160, Ap
C 23	44.2	7.1	465	4	Sequence 17442, A
C 24	44.2	7.1	465	4	Sequence 28116, A
C 25	43.8	7.0	476	4	Sequence 12361, A
C 26	43.8	7.0	1307	4	Sequence 1, Appli
C 27	43.2	6.9	1512	3	US-08-476-509B-1

C 28	43.2	6.9	1626	3	US-08-348-518C-1	Sequence 1, Appli
C 29	43.2	6.9	2208	3	US-08-539-205A-1	Sequence 1, Appli
C 30	43.2	6.9	2208	4	US-09-392-163A-1	Sequence 1, Appli
C 31	42.8	6.9	5115	3	US-08-348-518C-3	Sequence 3, Appli
C 32	42.8	6.9	5115	3	US-08-476-509B-3	Sequence 3, Appli
C 33	40	6.4	7218	1	US-08-232-463-14	Sequence 14, Appli
C 34	39.8	6.4	1463	4	US-09-270-767-15085	Sequence 15085, A
C 35	39.8	6.4	6396	4	US-09-620-312D-226	Sequence 226, App
C 36	39.2	6.3	4232	2	US-08-710-249-1	Sequence 1, Appli
C 37	39.2	6.3	4232	3	US-09-220-157A-1	Sequence 3, Appli
C 38	37.4	6.0	1005	2	US-08-844-312-3	Sequence 1, Appli
C 39	37.4	6.0	1290	2	US-08-844-312-1	Sequence 12556, A
C 40	36	5.8	39293	4	US-09-949-016-12556	Sequence 16078, A
C 41	36	5.8	39295	4	US-09-949-016-16078	Sequence 824, App
C 42	35.6	5.7	1074	4	US-09-134-000C-824	Sequence 2436, Ap
C 43	35.6	5.7	6347	4	US-09-949-016-2436	Sequence 14178, A
C 44	35.6	5.7	88268	4	US-09-949-016-14178	Sequence 5403, Ap
C 45	34.8	5.6	594	4	US-09-248-796A-5403	

ALIGNMENTS

RESULT 1
US-08-630-916A-45/c
; Sequence 45, Application US/08630916A
; Patent No. 6011137
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; APPLICANT: Kay, Brian K.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; TITLE OF INVENTION: POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 124
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: United States
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/630,916A
; FILING DATE: 03-APR-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-203
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2052 bases
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA
US-08-630-916A-45

Query Match 90.7%; Score 566.2; DB 3; Length 2052;
Best Local Similarity 97.3%; Pred. No. 5.3e-172;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

Qy 27 CCCTTTAGTCAGATGACTTCCCATGCGAGGATCTTTGAATCTTGTGTTCTTGTGTTA 86
|||
Db 1044 CCTTTAGTTACAGATGACTTCCCATGCGAGGATCTTTGAATGTTGTGTTCTTGTGTTA 985


```
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
US-09-070-060-2

Query Match      39.5%; Score 246.6; DB 2; Length 2559;
Best Local Similarity 64.3%; Pred. No. 7.3e-69;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

Qy 12 TAAGCAATTTGTGGACCCCTTTAGTNCAGATGACATCCCATTCGCGAGGATCTTTGAATGTT 71
Db 1418 TAGGCTATCTGAGTCCATGTTCTAGGCGAGATTTCTCTGCGGGATCTATATAGTG 1359

Qy 72 GTTGTCTTGTGTATGATCAAAAGTACCTTACACCTTCCAGAGTATATCTAAATTTCC 131
Db 1358 GTAGTCTTCTATTTGTGGTCCACAAAATATGGAATTCATCCCACTGTGAATCTCATTTCC 1299

Qy 132 CAGCTTCTGCGAGGGGTTCTTCAATCTGTAAGCCTTGAGTCTTGGATCTTCCCACTGG 191
Db 1298 CAACCTTCAGTAAGGGCTTTTCAATTAATGACCTTGAATCTTGGGGTCTTCCCATGT 1239

Qy 192 GTTGTCTTGTGTATGATCAAAAGTAAACCTGCTGTGTAATCCACTCTTTTTTTC 251
Db 1238 GTAATTCGTGTGTGTTGACGAAATATACTCTGCCATTGCTGTCTCTCTCTCC 1179

Qy 252 CAGCTGGTGGCAAGGTCCATAAGGTCATTTTCT-----GCAGCTAAC 296
Db 1178 CATCCAGTGGCAATGACCAAGAGGATCAAAATCTTTACTTTGTGATGTAGCAAAATAA 1119

Qy 297 ATTGAGCCGNAATAGAGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAATTTGTTTC 356
Db 1118 TCTTGATTTCCATAAATGAATCTCTGGTTAAACTGCTGCATTTGCTCTTGAAGCTGACTA 1059

Qy 357 CGCTGAGATTCGCACTGNTCAAATTCGGAAGATTCATGTTAGGCGCGCTGCCAGTT 416
Db 1058 CGCTGAGTGGCAATGTTTCACTAGTTCGGAAGATTCAGTGTGTGGCTCTGCCAGTT 999

Qy 417 GTTGTCTTGTGTATGATCCACATAATAAATCTCTCCATGTGTAGTTCGAGCATTTATGANCACA 536
Db 998 GTTGTCTTGTGAAATGGTCAACATAATAAATACGTCCTCCATGTTGTCAACCCGCGTTCC 939

Qy 477 CAACCTGGAGTAAAGTGTGTGCTCTCCCATGTGTAGTTCGAGCATTTATGANCACA 536
Db 938 CAGCAGGAGGTAGAGTGTGTCTATCCCATGTTGTTCTTTCTCAACATGATCTACA 879

Qy 537 TAATACGCTTACCATGAGGATCTTTTCTTTGTTCCACCCCTGATGGCAANGTTCTGT 596
Db 878 TAGTAAACTCGCCGCTGCTGCTCACTCTCTGCTCCCAACAGGTGSCAAGGAGCTTGA 819

Qy 597 GTT 599
Db 818 GTT 816

RESULT 4
US-09-357-746-2/c
; Sequence 2, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Fast-Seq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 2559
; TYPE: DNA
```

```
; ORGANISM: Homo sapiens
US-09-357-746-2

Query Match      39.5%; Score 246.6; DB 3; Length 2559;
Best Local Similarity 64.3%; Pred. No. 7.3e-69;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

Qy 12 TAAGCAATTTGTGGACCCCTTTAGTNCAGATGACATCCCATTCGCGAGGATCTTTGAATGTT 71
Db 1418 TAGGCTATCTGAGTCCATGTTCTAGGCGAGATTTTCTCTGCGGGATCTATATAGTG 1359

Qy 72 GTTGTCTTGTGTATGATCAAAAGTACCTTACACCTTCCAGAGTATATCTAAATTTCC 131
Db 1358 GTAGTCTTCTATTTGTGGTCCACAAAATATGGAATTCATCCCACTGTGAATCTCATTTCC 1299

Qy 132 CAGCTTCTGCGAGGGGTTCTTCAATCTGTAAGCCTTGAGTCTTGGATCTTCCCACTGG 191
Db 1298 CAACCTTCAGTAAGGGCTTTTCAATTAATGACCTTGAATCTTGGGGTCTTCCCATGT 1239

Qy 192 GTTGTCTTGTGTATGATCAAAAGTAAACCTGCTGTGTAATCCACTCTTTTTTTC 251
Db 1238 GTAATTCGTGTGTGTTGACGAAATATACTCTGCCATTGCTGTCTCTCTCTCC 1179

Qy 252 CAGCTGGTGGCAAGGTCCATAAGGTCATTTTCT-----GCAGCTAAC 296
Db 1178 CATCCAGTGGCAATGACCAAGAGGATCAAAATCTTTACTTTGTGATGTAGCAAAATAA 1119

Qy 297 ATTGAGCCGNAATAGAGTATCGTTGGTTAAACTGTTGCATAGCTCCCTGCAATTTGTTTC 356
Db 1118 TCTTGATTTCCATAAATGAATCTCTGGTTAAACTGCTGCATTTGCTCTTGAAGCTGACTA 1059

Qy 357 CGCTGAGATTCGCACTGNTCAAATTCGGAAGATTCATGTTAGGCGCGCTGCCAGTT 416
Db 1058 CGCTGAGTGGCAATGTTTCACTAGTTCGGAAGATTCAGTGTGTGGCTCTGCCAGTT 999

Qy 417 GTTGTCTTGTGTATGATCCACATAATAAATCTCTCCATGTGTAGTTCGAGCATTTATGANCACA 476
Db 998 GTTGTCTTGTGAAATGGTCAACATAATAAATACGTCCTCCATGTTGTCAACCCGCGTTCC 939

Qy 477 CAACCTGGAGTAAAGTGTGTGCTCTCCCATGTGTAGTTCGAGCATTTATGANCACA 536
Db 938 CAGCAGGAGGTAGAGTGTGTCTATCCCATGTTGTTCTTTCTCAACATGATCTACA 879

Qy 537 TAATACGCTTACCATGAGGATCTTTTCTTTGTTCCACCCCTGATGGCAANGTTCTGT 596
Db 878 TAGTAAACTCGCCGCTGCTGCTCACTCTCTGCTCCCAACAGGTGSCAAGGAGCTTGA 819

Qy 597 GTT 599
Db 818 GTT 816

RESULT 5
US-09-919-039-234/c
; Sequence 234, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR FILING DATE: 2002-09-09
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 234
; LENGTH: 2678
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 1989186CB1
```

US-09-919-039-234		Query Match		39.5%; Score 246.6; DB 4; Length 2678;	
Best Local Similarity 64.3%; Pred. No. 7.4e-69;		Best Local Similarity 64.3%; Pred. No. 1.1e-68;		DB 3; Length 5372;	
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;		Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;			
QY	12	TAAGCAATTTGTGGACCCCTTTATGNCAGATGACTCCCATTCGCGAGGATCTTTGAATGTT	71	12	TAAGCAATTTGTGGACCCCTTTATGNCAGATGACTCCCATTCGCGAGGATCTTTGAATGTT
DB	1190	TAGGCTATCTGAGGTCATTTCTAGGGCAGATTTTCTGTGCGGGGATCTATATAGGTG	1131	1567	TAGGCTATCTGAGGTCATTTCTAGGGCAGATTTTCTGTGCGGGGATCTATATAGGTG
QY	72	GTGTGTTCTGTGTATGATCAACAAGTACCTTACACCTTTCACGAGTATATCTAATTTCC	131	72	GTGTGTTCTGTGTATGATCAACAAGTACCTTACACCTTTCACGAGTATATCTAATTTCC
DB	1130	GTAGTTCCTTATTTGTGGTCCAAATATGGAATTCGATCCACTGTGAATCTCATTTCC	1071	1507	GTAGTTCCTTATTTGTGGTCCAAATATGGAATTCGATCCACTGTGAATCTCATTTCC
QY	132	CAGCCTTCTGGCAGGGGTTCTTCATTTCTGTAAAGCCTTGAGTTCTTGGATCTTCCCACTGG	191	132	CAGCCTTCTGGCAGGGGTTCTTCATTTCTGTAAAGCCTTGAGTTCTTGGATCTTCCCACTGG
DB	1070	CAACCTTCAGGTAAGGGCTTTTCATTTAATTAAGCCTTGAGTTCTTGGGCTTCCCATGT	1011	1447	CAACCTTCAGGTAAGGGCTTTTCATTTAATTAAGCCTTGAGTTCTTGGGCTTCCCATGT
QY	192	GTGTGTTTGTGTATGATCAACAAGTAAACCTGCTGTGTAATCCACTCTTTTTC	251	192	GTGTGTTTGTGTATGATCAACAAGTAAACCTGCTGTGTAATCCACTCTTTTTC
DB	1010	GTAAATTCGTGTGTGTGTTGACGAAATATATCTGCGCAATGCTGTCTCTCTCTCTCC	1328	1387	GTAAATTCGTGTGTGTGTTGACGAAATATATCTGCGCAATGCTGTCTCTCTCTCTCC
QY	252	CAGCCTGTGGCAAGGTCCTAAGGTCATTTTCT	296	252	CAGCCTGTGGCAAGGTCCTAAGGTCATTTTCT
DB	950	CATCCAGTGGCAATGGAACAGAGGATCAATTTCTTTACTTTGTGATGTAGCAATATA	891	1327	CATCCAGTGGCAATGGAACAGAGGATCAATTTCTTTACTTTGTGATGTAGCAATATA
QY	297	ATTGAAGCCGAATPAGAGGTATCGTTGTTAAACTGTGTGATGATGCTCCCTGCAATGGTTC	356	297	ATTGAAGCCGAATPAGAGGTATCGTTGTTAAACTGTGTGATGATGCTCCCTGCAATGGTTC
DB	890	TCITGATTTCCATAAATGAATCTCTGTTAAACTGTGCAATGCTCTTGAAGCTGACTA	831	1267	TCITGATTTCCATAAATGAATCTCTGTTAAACTGTGCAATGCTCTTGAAGCTGACTA
QY	357	CGCTGAGATGCGCACTGNTCAAAATTCGACAGATTCATGATGAGCGCTGCCACGTT	416	357	CGCTGAGATGCGCACTGNTCAAAATTCGACAGATTCATGATGAGCGCTGCCACGTT
DB	830	CGCTGAGTGGCAATGTTTATAGTTTCCGACGAGATTCAGTGTGCGCTCTGCCACGTT	771	1207	CGCTGAGTGGCAATGTTTATAGTTTCCGACGAGATTCAGTGTGCGCTCTGCCACGTT
QY	417	GTGTGTTCTGTGTATGATCCACATATAAATCTCTACGATCATCACTCTTCTTCC	476	417	GTGTGTTCTGTGTATGATCCACATATAAATCTCTACGATCATCACTCTTCTTCC
DB	770	GTGTGTTCTGTGAATGGTCAACATAAATAGTCCCATGTTGTCAACCGCGCTTC	711	1147	GTGTGTTCTGTGAATGGTCAACATAAATAGTCCCATGTTGTCAACCGCGCTTC
QY	477	CAACCTGGAGGTAAGGTTGTGTCTCCCATGTGTAGTTCGAGCATATGANCACA	536	477	CAACCTGGAGGTAAGGTTGTGTCTCCCATGTGTAGTTCGAGCATATGANCACA
DB	710	CAGCAGGAGGTAGAGGTTCTGTCTATCCCATGTTTCTTCTTCTCAACATGATCTACA	651	1087	CAGCAGGAGGTAGAGGTTCTGTCTATCCCATGTTTCTTCTTCTCAACATGATCTACA
QY	537	TAATACGCTCTACCATGAGGATCTTTTCTTGTTCACCCCTGTAGTGGCAANGTTCTGT	596	537	TAATACGCTCTACCATGAGGATCTTTTCTTGTTCACCCCTGTAGTGGCAANGTTCTGT
DB	650	TAGTAACCTCGCCGTCGTGTCACCTCTCTCTCCCAACAGGTGGCAAGGAGCTTGA	591	1027	TAGTAACCTCGCCGTCGTGTCACCTCTCTCTCCCAACAGGTGGCAAGGAGCTTGA
QY	597	GTT 599		597	GTT 599
DB	590	GTT 588		967	GTT 965

RESULT 7
US-09-070-060-1/c
; Sequence 1, Application US/09070060
; Patent No. 5976849
; GENERAL INFORMATION:
; APPLICANT: Husted, Carolyn M.
; APPLICANT: Ghildyal, Namit
; TITLE OF INVENTION: Human E3 Ubiquitin Protein
; TITLE OF INVENTION: Ligase
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZENECA Pharmaceuticals, Inc.
; STREET: 1800 Concord Pike
; CITY: Wilmington
; STATE: DE
; COUNTRY: USA
; ZIP: 19850-5437
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/070,060

RESULT 6
US-09-357-746-1/c
; Sequence 1, Application US/09357746
; Patent No. 6087122
; GENERAL INFORMATION:
; APPLICANT: ZENECA Limited
; TITLE OF INVENTION: HUMAN E3 UBIQUITIN PROTEIN LIGASE
; FILE REFERENCE: PHM.70312.N1
; CURRENT APPLICATION NUMBER: US/09/357,746
; CURRENT FILING DATE: 1999-07-21
; EARLIER APPLICATION NUMBER: US No. 6087122 60/073,839
; EARLIER FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: US No. 608712209/070,060
; EARLIER FILING DATE: 1998-04-30
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 1
; LENGTH: 5372
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-357-746-1


```
; APPLICATION NUMBER: US/08/895,601
; FILING DATE: 16-JUL-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: MIV-096.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-832-1000
; TELEFAX: 617-832-7000
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2790 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..2782
; US-08-895-601-2

Query Match      8.7%; Score 54; DB 3; Length 2790;
Best Local Similarity 48.5%; Pred. No. 1e-06;
Matches 178; Conservative 0; Mismatches 183; Indels 6; Gaps 1;

QY 173 TCTGGATCTCCCACTGGGTTGTTTGTATGATTCACAAAGTAAACCCCTGTCTGT 232
DB 1441 TCTTGGATCTTCCAGGTGGTTTGTAGTTGTGTCATTAAGAAAGGCTCCCAT 1382
QY 233 TGAATCCACTCTTTTCCAGCTGTGTGCAAGTCCATATAGGTCATTTCTGAGC 292
DB 1381 TGGTGCATGCCGACTTCCAGGCTTTAGGAAGAAATCCCTTGC-----TCAATTTTCA 1328
QY 293 TAACATTGAAGCCGAATAGAGGTATCGTTGGTTAAACCTGTGCATAGCTCCCTGCAATTG 352
DB 1327 TGGCTGGTCACTGTGGCTGAATCACTGTGGAGGCTTGATGATGAGGCTGCGA 1268
QY 353 GTTCCCGTGAGATGCGCACTGNTCAAAATTCGACAGATTCATCGTAGGCGCTGCCA 412
DB 1267 ACTCTGGCTTGAGGTGAGTCACTGCTCCACTGTGGCTGTACAGTGGGCTTTGTCCA 1208
QY 413 CGTTGTTGTTCTGTTATGATCCACATTAATACTCTCTACGATCACTCTTCT 472
DB 1207 AGTAGTGGTTCTGGAATTGGATCATATAATATGATCTTCTCTTTTCATCTTGTTC 1148
QY 473 TTCCCAACCTGGAGGTAAGGTTGTGCTCTCCCATGTGGTAGTTCGAGCATTATGANC 532
DB 1147 TTCCCAACCTGGTGGTAATCCAGATGAGTACCAAGACACAGGAAGTAGGTGTTTC 1088
QY 533 CACATAA 539
DB 1087 CTCAAAA 1081

RESULT 12
US-09-949-016-1035/c
; Sequence 1035, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1035
; LENGTH: 2264
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-1035

Query Match      8.6%; Score 53.8; DB 4; Length 2264;
Best Local Similarity 54.9%; Pred. No. 1e-06;
Matches 128; Conservative 0; Mismatches 102; Indels 3; Gaps 1;

QY 57 GGATCTTTGAATGTTGTTCTTGTGTTATGATCAACAAGTACCTTACACCTTCACGA 116
DB 388 GGGTCCAAGTAGGTGTTCTTTATTATATGTTCAACAAAAACACTTGTCCGTTCTCA 329
QY 117 GTATATCTAATTTCCAGCCTTCTGGCAGGGGTTCTTCAITCTCTAGACCTT---GAGTT 173
DB 328 TCAGTTTCTTGTTCCTCATCGTATGCAAAATCTCTGCCACTCGTTTCTTTTCCAGTT 269
QY 174 CTTGGATCTTCCCACTGGGTTGTTTGTGTTATGATCAAAAGTAAACCTGTCTGTT 233
DB 268 TTTGGATGTTCCCACTGAGTCTTCTCTCGGTGTGATTGGCGTAGTAAACCCAGCCGTC 209
QY 234 GAATCCACTCTTTTTCAGCCTGTGGCAAGGTCCTAAGGTCATTTTC 286
DB 208 TTGGTGGTTCTCTCTCTCCAGCCGCGAGGAGCTGCTCCTCACTGTCCTGTC 156

RESULT 13
US-08-539-205A-3/c
; Sequence 3, Application US/08539205A
; Patent No. 6001619
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/539,205A
; FILING DATE: 04-OCT-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..2535
; US-08-539-205A-3

Query Match      8.6%; Score 53.4; DB 3; Length 2848;
```

US-09-248-796A-4083/c
; Sequence 4083, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; PRIOR FILING DATE: 1999-02-12
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-02-13
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 4083
; LENGTH: 267
; TYPE: DNA
; ORGANISM: Candida albicans
US-09-248-796A-4083

Query Match 8.4%; Score 52.2; DB 4; Length 267;
Best Local Similarity 74.2%; Pred. No. 1.1e-06;
Matches 66; Conservative 0; Mismatches 23; Indels 0; Gaps 0;

QY 402 GGCCGCTGCCACGTTGTTTCTGTGTATGATCCACATAATAAATCTCTTACGATCA 461
DB 191 GGACGTTGCCAAGTTGTAGTTCTAGAAATTTGATCAACATATAATAGTTCTACCAAAATTA 132

QY 462 TCAACTCTTCTTTCCCAACCTGGAGGTAA 490
DB 131 TCAGTTCTACGTTCCCAACCCAGGTGGTAA 103

Search completed: June 30, 2005, 05:41:07
Job time : 162 secs

US-09-392-163A-3/c
; Sequence 3, Application US/09392163A
; Patent No. 6503742
; GENERAL INFORMATION:
; APPLICANT: Beach, David H.
; APPLICANT: Caligiuri, Maureen
; APPLICANT: Nefsky, Bradley
; TITLE OF INVENTION: Ubiquitin Ligases, and Uses Related Thereto
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY, HOAG & ELIOT LLP
; STREET: One Post Office Square
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109-2170
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/392,163A
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/539,205
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Vincent, Matthew P.
; REGISTRATION NUMBER: 36,709
; REFERENCE/DOCKET NUMBER: CSV-005.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 832-1000
; TELEFAX: (617) 832-7000
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2848 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 238..2535
US-09-392-163A-3

Query Match 8.6%; Score 53.4; DB 4; Length 2848;
Best Local Similarity 69.9%; Pred. No. 1.6e-06;
Matches 72; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 387 ACAGATTCATGCTAGGCGCTGCCACGTTGTTCTGTGTATGATCCACATAATAA 446
DB 959 ACAGAACTTAGTTGGGGCAATCCACGTTGCTTCTGTATGATGATCTACATAATAA 900

QY 447 ACTCCTCTACGATCATCAACTCTTCTTTCCCAACCTGGAGGTA 489
DB 899 GTTCGGCCAAAGATTATCGGTAGTCTTTCCCATCCAGGAGGAA 857

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 03:09:44 ; Search time 3088 Seconds
(without alignments)
9791.468 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtntaagcattt.....tccagactgctgccgnaca 624

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hgt.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	566.2	90.7	2052	6	AR162499 Sequence
c 2	566.2	90.7	2052	6	AX675520 Sequence
c 3	566.2	90.7	2052	9	HSU96113 Sequence
c 4	566.2	90.7	2827	6	AX053140 Sequence
c 5	566.2	90.7	3382	9	HSM801707 Sequence
c 6	566.2	90.7	3479	9	AY043361 Homo sapi
c 7	566.2	90.7	3599	9	AY345857 Homo sapi
c 8	566.2	90.7	3855	9	BC036065 Homo sapi
c 9	552.2	90.7	3855	6	AX053142 Sequence
c 10	510.8	81.9	49548	9	AC016962 Homo sapi
c 11	431.2	69.1	1794	5	AF412122 Gallus ga
c 12	367.8	58.9	3501	5	BC076842 Xenopus l
c 13	355.4	57.0	130034	2	AC101726 Mus muscu
c 14	355.4	57.0	135660	2	AC103927 Mus muscu
c 15	355.4	57.0	189541	2	AC101785 Mus muscu
c 16	355.4	57.0	233321	10	AL732546 Mouse DNA
c 17	347	55.6	254053	2	AC109017 Rattus no
c 18	255.2	40.9	1867	5	CR386634 Gallus ga
c 19	249.6	40.0	2565	10	AY600518 Rattus no

c 20	249.2	39.9	5122	10	AF037454 Mus muscu
c 21	249.2	39.9	5163	10	BC062934 Mus muscu
c 22	249.2	39.9	5182	10	BC064678 Mus muscu
c 23	246.6	39.5	2377	6	AX277586 Sequence
c 24	246.6	39.5	2377	6	AX675519 Sequence
c 25	246.6	39.5	2377	6	AX774812 Sequence
c 26	246.6	39.5	2377	9	AF038564 Homo sapi
c 27	246.6	39.5	2559	6	AR083345 Sequence
c 28	246.6	39.5	2559	6	AR102842 Sequence
c 29	246.6	39.5	2559	6	BD124048 Protein.
c 30	246.6	39.5	2678	6	AR531671 Sequence
c 31	246.6	39.5	2869	9	AB056663 Homo sapi
c 32	246.6	39.5	2970	9	AF095745 Homo sapi
c 33	246.6	39.5	3745	9	BC011571 Homo sapi
c 34	246.6	39.5	5372	6	AR102841 Sequence
c 35	246.6	39.5	5372	6	BD124047 Protein.
c 36	246.2	39.5	5359	6	AR083344 Sequence
c 37	233.8	37.5	180976	9	AL513321 Human DNA
c 38	213.2	34.2	2910	10	BC039921 Mus muscu
c 39	213.2	34.2	4319	10	BC048184 Mus muscu
c 40	211.6	33.9	4442	9	BC013645 Homo sapi
c 41	211.6	33.9	4500	9	BC064531 Homo sapi
c 42	203.6	32.6	3475	6	AR162506 Sequence
c 43	203.6	32.6	3475	6	AX675521 Sequence
c 44	203.6	32.6	3475	9	HSU96114 Homo sapien
c 45	201.2	32.2	2394	6	CQ727501 Sequence

ALIGNMENTS

RESULT 1
AR162499/c
LOCUS AR162499 2052 bp DNA linear PAT 17-OCT-2001
DEFINITION Sequence 3 from patent US 6258601.
ACCESSION AR162499
VERSION AR162499.1 GI:16229714
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2052)
AUTHORS Monia,B.P. and Cowser,L.M.
TITLE Antisense modulation of ubiquitin protein ligase expression
JOURNAL Patent: US 6258601-A 3 10-JUL-2001;
FEATURES
source Location/Qualifiers
1..2052
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN

Query Match	90.7%	Score	566.2;	DB	6;	Length	2052;
Best Local Similarity	97.3%	Pred. No.	2.1e-144;				
Matches	582;	Conservative	0;	Mismatches	15;	Indels	1;
Gaps	1;						
Qy	27	CCCTTAGTACAGTACCTCCATCCGAGGATCTTTGAATGTGTGTCTGTGTTA	86				
Db	1044	CCCTTAGTACAGTACCTCCATCCGAGGATCTTTGAATGTGTGTCTGTGTTA	985				
Qy	87	TGATCAACAAGTACCTTACACCTTCCAGGATATCTAATTTCCAGCCTTCTGGCAGG	146				
Db	984	TGATCAACAAGTACCTTACACCTTCCAGGATATCTAATTTCCAGCCTTCTGGCAGG	925				
Qy	147	GGTCTTCATCTGTGAAGCCTTCAGTCTTGGATCTTCCACATCTGGGTGTGTGTGTTA	206				
Db	924	GGTCTTCATCTGTGAAGCCTTCAGTCTTGGATCTTCCACATCTGGGTGTGTGTGTTA	865				
Qy	207	TGATTCACAAGTAAACCTCTGTGTGAATCACTCTTTTCCAGCCTGTGGCAGG	266				
Db	864	TGATTCACAAGTAAACCTCTGTGTGAATCACTCTTTTCCAGCCTGTGGCAGG	805				
Qy	267	GGTCCATAGGGTTCATTTCTGAGCTACATTTGAAGCCGAATAGAGTATCGTTGTTA	326				

Db 804 GGTCCATAGAGGTCATTTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

Qy 327 AACTGTTGATAGTCCCTGCAATTTGGTTCCTGAGATTGCGACTGNTCAAAATTCGG 386

Db 744 AACTGTTGATAGTCCCTGCAATTTGGTTCCTGAGATTGCGACTGNTCAAAATTCGG 685

Qy 387 ACAGATTCATAGTGGCGCTGCCACGTTGTTCTGTTGTTAAGATCCACATAATAA 446

Db 684 ACAGATTCATAGTGGCGCTGCCACGTTGTTCTGTTGTTAAGATCCACATAATAA 625

Qy 447 ACTCCTCTACGATCATCAACTCTCTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 506

Db 624 ACTCCTCTACGATCATCAACTCTCTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 565

Qy 507 CATGTGGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 566

Db 564 CATGTGGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505

Qy 567 TGTTCCTCCACCTGATGGCAANGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCCGNACA 624

Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCCGTACA 448

RESULT 2

AX675520/c

LOCUS AX675520 2052 bp DNA linear PAT 27-MAR-2003

DEFINITION Sequence 3 from Patent WO02090549.

ACCESSION AX675520

VERSION AX675520.1 GI:29333538

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Greener, T., Moskowitz, H., Reiss, Y., and Alroy, I. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2052)

AUTHORS Kay, B.K. and Fowlkes, D.M.

TITLE Identification of novel human WW domain-containing proteins by cloning of ligand targets

JOURNAL J. Biol. Chem. 272 (23), 14611-14616 (1997)

MEDLINE 97313427

PUBMED 9169421

REFERENCE 2 (bases 1 to 2052)

AUTHORS Pirozzi, G. and Uveges, A.

TITLE Direct Submission

JOURNAL Submitted (02-APR-1997) CytoGen Corp., 201 College Road East, Princeton, NJ 08540, USA

FEATURES

source

1..2052

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

ORIGIN

Query Match 90.7%; Score 566.2; DB 6; Length 2052;

Best Local Similarity 97.3%; Pred. No. 2.1e-144;

Matches 592; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

Qy 27 CCCTTTAGTNCAGATGACTTCCCATGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86

Db 1044 CCTTTAGTTACAGATGACTTCCCATGCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 985

Qy 87 TGAATCAACAAAGTACCTTACACTTCCAGAGTATCTAAATTCACGCTTCTGGCAGG 146

Db 984 TGAATCAACAAAGTACCTTACACTTCCAGAGTATCTAAATTCACGCTTCTGGCAGG 925

Qy 147 GGTTCCTTCATCTGTAAGCCTTGAGTTCTGATCTTCCACTGCTGTTGTTGTTGTTA 206

Db 924 GGTTCCTTCATCTGTAAGCCTTGAGTTCTGATCTTCCACTGCTGTTGTTGTTGTTA 865

Qy 207 TGAATCAACAAAGTAAACCTTCTGTTGAATCCACTCTTTTTCACGCTGCTGGCAGG 266

Db 864 TGAATCAACAAAGTAAACCTTCTGTTGAATCCACTCTTTTTCACGCTGCTGGCAGG 805

Qy 267 GGTCCATAGAGGTCATTTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326

Db 804 GGTCCATAGAGGTCATTTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

Qy 327 AACTGTTGATAGTCCCTGCAATTTGGTTCCTGAGATTGCGACTGNTCAAAATTCGG 386

Db 744 AACTGTTGATAGTCCCTGCAATTTGGTTCCTGAGATTGCGACTGNTCAAAATTCGG 685

Qy 387 ACAGATTCATAGTGGCGCTGCCACGTTGTTCTGTTGTTAAGATCCACATAATAA 446

Db 684 ACAGATTCATAGTGGCGCTGCCACGTTGTTCTGTTGTTAAGATCCACATAATAA 625

Qy 447 ACTCCTCTACGATCATCAACTCTCTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 506

Db 624 ACTCCTCTACGATCATCAACTCTCTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 565

Qy 507 CATGTGGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 566

Db 564 CATGTGGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505

Qy 567 TGTTCCTCCACCTGATGGCAANGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCCGNACA 624

Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATTCCTCCAGACTGCTGCCGTACA 448

RESULT 3

HSU96113/c

LOCUS HSU96113 2052 bp mRNA linear PRI 29-MAY-1997

DEFINITION Homo sapiens Nedd-4-like ubiquitin-protein ligase WWP1 mRNA, partial cds.

ACCESSION U96113

VERSION U96113.1 GI:2072500

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2052)

AUTHORS Pirozzi, G., McConnell, S.J., Uveges, A.J., Carter, J.M., Sparks, A.B., Kay, B.K. and Fowlkes, D.M.

TITLE Identification of novel human WW domain-containing proteins by cloning of ligand targets

JOURNAL J. Biol. Chem. 272 (23), 14611-14616 (1997)

MEDLINE 97313427

PUBMED 9169421

REFERENCE 2 (bases 1 to 2052)

AUTHORS Pirozzi, G. and Uveges, A.

TITLE Direct Submission

JOURNAL Submitted (02-APR-1997) CytoGen Corp., 201 College Road East, Princeton, NJ 08540, USA

FEATURES

source

1..2052

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

<1..>2052

/note="Nedd-4-like ubiquitin-protein ligase; WW domain-containing protein"

/codon_start=2

/product="WWP1"

/protein_id="AAC51324.1"

/db_xref="GI:2072501"

/translation="TNHVPTSLVQNSCCSYVNGDTPSSPSQVAARPNTPAKPL ASPADLTNGSSSPAPDNTASVTGTPVSEENALSPNCTSTTVDPVQBLTSSSE ANECIPSTSAELESEARSILEPDTNSRSSSAPEAKSRQPDGMDPVQSGNANTE TLPSEQRNDPHGRYTYVDHNTTRTTTWERPQLPPGWERRVDDRRVYVDHNTTTR VSDTRVYFVHNKTKTQWEDPRKQLQNEBPELPEGWETRYREGVRYFVDHNTTTR FKDPNGKSVTVTGGPQIAYBERGFRWLKLAHFRYLCQSNALPSHVKLINVSQTLPEDSF QOJMALPKVDLRRRLVYIFRBERGLDYGGLAREWFFLLSHEVLNPMYCLFETAGKNY CLOINPASTINPDHLSVFCPIGRFIAMALFHGKFIIDTGFSLDPLKGLSNILVTENKDE SITDEFNSIWIETDNNIRECGLMEYFSDVMEILGKVTSHDLKGLSNILVTENKDE YIGLMTWFRSGVQSQTKAFLDGFNEVVPLOWLYFDEKELEVMCLGMEVLDADMQ RNTVRYHTRNSKQIIFWFOFVKTENVRMELLQFVTGTCRLPLGFGAELMGSGNPR NS"

ORIGIN

Query Match 90.7%; Score 566.2; DB 9; Length 2052;

Best Local Similarity 97.3%; Pred. No. 2.1e-144;

Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

```
QY 27 CCCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTTGAATGTTGTTGTTCTTTGTGTTA 86
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTTGAATGTTGTTGTTCTTTGTGTTA 985
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 87 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 146
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 984 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 925
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 147 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 924 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 864
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCACAGCCTGGTGGCAAA 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 924 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 865
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCACAGCCTGGTGGCAAA 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 864 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCACAGCCTGGTGGCAAA 805
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 326
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 804 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 745
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGCG 386
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 744 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGCG 685
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 387 ACAGATTCCATGTTAGCCCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATAA 446
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 684 ACAGATTCCATGTTAGCCCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATAA 625
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCACTGGAGTAAAGTTGTTGTTCTCTCC 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 624 ACTCCTCTACGATCATCACTCTTCTTCCCACTGGAGTAAAGTTGTTGTTCTCTCC 565
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 507 CATGTGTTAGTTCGAGCATTTATGATCCACATAATACGCTCTACCATGAGGATCTTTTCTT 566
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 564 CATGTGTTAGTTCGAGCATTTATGATCCACATAATAGTTCTACCATGAGGATCTTTTCTT 505
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 567 TGTTCCACCCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 504 TGTTCCCA-CCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 448
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 4
AX053140/c
LOCUS AX053140 2827 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 2 from Patent WO0073329.
ACCESSION AX053140
VERSION AX053140.1 GI:12227505
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
REFERENCE Baron, M.
AUTHORS Drosophila sp.
TITLE Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
JOURNAL Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
FEATURES
source
1. .2827
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"

Query Match 90.7%; Score 566.2; DB 6; Length 2827;
Best Local Similarity 97.3%; Pred. No. 2.1e-144;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTTGAATGTTGTTGTTCTTTGTGTTA 86
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1043 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTTGAATGTTGTTGTTCTTTGTGTTA 984
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
QY 87 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 146
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 983 TGATCAACAAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACAGCCTTCTGGCAGG 924
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 147 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 206
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 923 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCTTCCACCTGGGTGTTTGTGTTA 864
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCACAGCCTGGTGGCAAA 266
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 863 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTTCCACAGCCTGGTGGCAAA 804
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 267 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCCCAATAGAGGATTCGTTGTTA 326
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 803 GGTTCCTTCATTCTGTAAAGCCCTTCAGTTCCTTGGATCCCAATAGAGGATTCGTTGTTA 744
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGCG 386
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 743 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTCGCG 684
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 387 ACAGATTCCATGTTAGCCCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATAA 446
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 683 ACAGATTCCATGTTAGCCCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATAA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCACTGGAGTAAAGTTGTTGTTCTCTCC 506
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 623 ACTCCTCTACGATCATCACTCTTCTTCCCACTGGAGTAAAGTTGTTGTTCTCTCC 564
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 507 CATGTGTTAGTTCGAGCATTTATGATCCACATAATAGTTCTACCATGAGGATCTTTTCTT 566
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 563 CATGTGTTAGTTCGAGCATTTATGATCCACATAATAGTTCTACCATGAGGATCTTTTCTT 504
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 567 TGTTCCACCCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 503 TGTTCCCA-CCCTGATGGCAAGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 447
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 5
HSM801707/c
LOCUS HSM801707 3382 bp mRNA linear PRI 20-MAR-2002
DEFINITION Homo sapiens mRNA; cDNA DKFZp434D2111 (from clone DKFZp434D2111).
ACCESSION AL136739
VERSION AL136739.1 GI:12052996
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Duesterhoeft, A., Lauber, J., Mewes, H.W., Weil, B. and Wiemann, S.
AUTHORS 1 (bases 1 to 3382)
TITLE Direct Submission
JOURNAL Submitted (12-MAR-2002) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
COMMENT Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
This clone (DKFZp434D2111) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at http://mips.gsf.de/proj/cDNA/.

FEATURES
Location/Qualifiers
1. .3382
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="RZPD:DKFZp434D2111"
/db_xref="taxon:9606"
/chromosome="3"
/clone="DKFZp434D2111"
/tissue_type="testis"
/clone_lib="434 (synonym: htes3). Vector pSport1; host
```

```

DH10B; sites NotI + SalI"
/dev_stage="adult"
1. .3382
/gene="DKFZp434D2111"
1..2779
/gene="DKFZp434D2111"
/note="similarity to Nedd-4-like ubiquitin-protein ligase
(Homo sapiens)"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB66673.1"
/db_xref="GI:12052397"
/db_xref="GOA:Q9H0M0"
/db_xref="UniProt/TrEMBL:Q9H0M0"
/translation="NATASPRSDTNNHSGRLQLQVTVSSAKLKKKKWFGTAIYEV
VVDGETTKAKSSSNPKWDQLTVNVPQTTFQVWSHRTLKADALLGKATIDLK
QALLHNRLKRVKEQLKLSLENGKIAQTGLTVLGLVIEQENITNCSSPTIEI
QENGDAHENGEPASRTTARLAVEGTNGIDNHVPTSTLVONSCCSVVNGDNTSPSP
QVAAPKNTPAKPLASEPADDTVNGESSFAPTDNASVTGTPVSEENALSPNCTST
TVEDPPVQBLTSSNNECIPSTSAESEAARSILEPDTNSRSSSAEAAKSRPDG
CMDPVRQSGNANTETLPSGWQRKDPHGRTYYVDHNTKTTTWERPQPLPGWERRV
DRRVVYVDHNTRTTTQWPTMESVNFQOSQORNLQGMQOQNRLYSASMLAA
ENDPYGLPGEKRVDSITRVYFVNHTKTTQWEDPRTQGLQNEEPLPEGWEIYTR
EGRVYVDHNTRTTTFKDPNGKSVTKGQPOIAYERGPRKLAHFRYLCQSNALPFSH
VKINVSQRTLFEDSFQOIMALKAPYDLRLRYLVIERGEGLDYGLAREWFFLLSHVL
NPMYCLFEYAGKNYCLQINPASTINPDHLSYFCFGRFIAMALPHKGFIDTGFSLPF
YKMLSKLTIKDLESIDTFYNSLIWIKDNNEIECGLMEYFSVMEILLKVTSHDLK
LGSGNLTVEENKDEYIGLMEWFRSGVQOEKAPLDFGNFVPLQWQVMEILLKVTSHDLK
VLMCGMQEVDLADWQNTVYRHRNSKQIIFWQFVKETDNEVRMLQFVTGRL
PLGGFAELMGSNGPQKFCIEKVGKDTLWLPRESHTCFNRLDLPYKSYEQLEKLLFAIE
ETEGFQGE"
3346..3351
/polyA_signal
3362
/gene="DKFZp434D2111"
/polyA_site
ORIGIN
Query Match 90.78; Score 566.2; DB 9; Length 3382;
Best Local Similarity 97.34; Pred. No. 2e-144; Indels 1; Gaps 1;
Matches 582; Conservative 0; Mismatches 15;
Qy 27 CCCTTTAGTCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 86
Db 1620 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTA 1561
Qy 87 TGATCAACAAAGTACCTTACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAG 146
Db 1560 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAG 1501
Qy 147 GGTCTTCATCTCTGAGCCCTTGAGTCTTGGATCTTTCCACTGGGTGTTTGTGTTA 206
Db 1500 GGTCTTCATCTCTGAGCCCTTGAGTCTTGGATCTTTCCACTGGGTGTTTGTGTTA 1441
Qy 207 TGATTCACAAAGTAAACCTGCTGTTGAAPCCACTCTTTTTTCCAGCCTGTTGCGAAA 266
Db 1440 TGATTCACAAAGTAAACCTGCTGTTGAATCCACTCTTTTTTCCAGCCTGTTGCGAAA 1381
Qy 267 GGTCAATAGGTCATTTTTCGACGCTAACATTCGACCCGAATAGAGTATCGTTGTTA 326
Db 1380 GGTCAATAGGTCATTTTTCGACGCTAACATTCGACCCGAATAGAGTATCGTTGTTA 1321
Qy 327 AACTGTTGCATAGCTCCCTGCAATGGTTCCGCTGAGATTGCCACTGNTCAAAATTNCGG 386
Db 1320 AACTGTTGCATAGCTCCCTGCAATGGTTCCGCTGAGATTGCCACTGNTCAAAATTNCGG 1261
Qy 387 ACAGATTCATGTTAGGCGCGTGCACCTGTTGTTGTTGTTGTTATGATCAACATAATA 446
Db 1260 ACAGATTCATGTTAGGCGCGTGCACCTGTTGTTGTTGTTGTTATGATCAACATAATA 1201
Qy 447 ACTCCCTCTACCATCATCACTCTTCTTCCCAACTCGAGGTAAAGGTTGTTGTTCTCTCC 506
Db 1200 ACTCTCTCTACCATCATCACTCTTCTTCCCAACTCGAGGTAAAGGTTGTTGTTCTCTCC 1141
```

```

507 CATGTGTTAGTTCGAGCATTATGANCACATAATACGCTCTACCATGAGGATCTTTCTT 566
1140 CATGTGTTAGTTCGAGTATTATGATCCACATAATAGGTTCTACCATGAGGATCTTTCTT 1081
567 TGTTCACACCCCTGATGGCAANGNTTCTGTGTGGCATTCACGACATGCTGCCGNACA 624
1080 TGTTCACCA-CCCTGATGCAAGGTTTCTGTGTGGCATTCACGACATGCTGCCGTACA 1024

RESULT 6
AY043361/c 3479 bp mRNA linear PRI 01-SEP-2001
LOCUS Homo sapiens WW domain-containing protein 1 mRNA, complete cds.
DEFINITION
ACCESSION AY043361
VERSION AY043361.1 GI:15419010
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3479)
AUTHORS Flasa,M., Aslam,H., Poulson,R., Hanby,A., Gorman,P., Roylance,R.,
Nolan,T., Brady,G., Canfield,A. and Baron,M.
WP1, a human homolog of Drosophila suppressor of deltex
TITLE Journal
REFERENCE 2 (bases 1 to 3479)
AUTHORS Flasa,M., Aslam,H., Poulson,R., Hanby,A., Gorman,P., Roylance,R.,
Nolan,T., Brady,G., Canfield,A. and Baron,M.
Direct Submission
JOURNAL Submitted (03-JUL-2001) Biological Sciences, University of
Manchester, Stoford Building, Oxford Rd., Manchester M13 9PT, UK
FEATURES
Location/Qualifiers
1..3479
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8q21"
86..2854
/note="WWP1; Nedd4 family E3 ubiquitin ligase; suppressor
of deltex related protein 1 (hSDRP1)"
/codon_start=1
/product="WW domain-containing protein 1"
/protein_id="AAK94668.1"
/db_xref="GI:15419011"
/translation="NATASPRSDTNNHSGRLQLQVTVSSAKLKKKKWFGTAIYEV
VVDGETTKAKSSSNPKWDQLTVNVPQTTFQVWSHRTLKADALLGKATIDLK
QALLHNRLKRVKEQLKLSLENGKIAQTGLTVLGLVIEQENITNCSSPTIEI
QENGDAHENGEPASRTTARLAVEGTNGIDNHVPTSTLVONSCCSVVNGDNTSPSP
QVAAPKNTPAKPLASEPADDTVNGESSFAPTDNASVTGTPVSEENALSPNCTST
TVEDPPVQBLTSSNNECIPSTSAESEAARSILEPDTNSRSSSAEAAKSRPDG
CMDPVRQSGNANTETLPSGWQRKDPHGRTYYVDHNTKTTTWERPQPLPGWERRV
DRRVVYVDHNTRTTTQWPTMESVNFQOSQORNLQGMQOQNRLYSASMLAA
ENDPYGLPGEKRVDSITRVYFVNHTKTTQWEDPRTQGLQNEEPLPEGWEIYTR
EGRVYVDHNTRTTTFKDPNGKSVTKGQPOIAYERGPRKLAHFRYLCQSNALPFSH
VKINVSQRTLFEDSFQOIMALKAPYDLRLRYLVIERGEGLDYGLAREWFFLLSHVL
NPMYCLFEYAGKNYCLQINPASTINPDHLSYFCFGRFIAMALPHKGFIDTGFSLPF
YKMLSKLTIKDLESIDTFYNSLIWIKDNNEIECGLMEYFSVMEILLKVTSHDLK
LGSGNLTVEENKDEYIGLMEWFRSGVQOEKAPLDFGNFVPLQWQVMEILLKVTSHDLK
VLMCGMQEVDLADWQNTVYRHRNSKQIIFWQFVKETDNEVRMLQFVTGRL
PLGGFAELMGSNGPQKFCIEKVGKDTLWLPRESHTCFNRLDLPYKSYEQLEKLLFAIE
ETEGFQGE"
137..508
/misc_feature
/note="Region: C2 domain"
1130..1222
/misc_feature
/note="Region: WW domain"
1223..1333
/misc_feature
/note="Region: WW domain"
1451..1543
/misc_feature
/note="Region: WW domain"
1571..1666
/misc_feature
/note="Region: WW domain"
1787..2851
/misc_feature
```

ORIGIN	/note="Region: HECT domain"		source	1. .3599
Query Match	90.7%;	Score 566.2; DB 9; Length 3479;	/organism="Homo sapiens"	
Best Local Similarity	97.3%;	Pred. No. 2e-144;	/mol_type="mRNA"	
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;		/db_xref="taxon:9606"	
			315_3083	
			/note="Tiul1; E3 ubiquitin ligase"	
			/codon_start=1	
			/product="TGIF-interacting ubiquitin ligase 1"	
			/protein_id="AAQ22764.1"	
			/db_xref="GI:33590476"	
			/translation="MATASPRSDTSNNHSRLQLOVTVSSAKLKRKNMFGTAIYEV	
			VIDGEITKTKASSSSNPWDQOLTVNTPQTLEQVWSHRTLKADALLKXKATIDLK	
			QALLIHNRLKRVKSLKSLNKGIAQBELTVVDGLVLIQSEINLTCSSPTTLEI	
			QENGDAHNGEPSEATLAVENGNGIDNHVPTSTLVQNSCCSVVSEVNDTPSPS	
			QVAAPKPTAPKPLASPADDTVNGESSFAPTDNASVTGPVSEVNSLNCSTST	
			TWDPVQIILTSSENNECI PSTSALESEARSILEPDTNSRSRFAEAKRQPDV	
			CMDVPQOSGNANTETLPSGMBQRKDPHGRTVVDHNTRTTWERPQPLPGWERRVD	
			DRRVYVVDHNTRTTWRPTWESVRNPEQWQSQRLQGMQFQRIYLSAMLA	
			ENDPGLPPEWEKRVDSIDRVYFNHNTKTQWEDPRTQGLQNEPFLPGWEIRYTR	
			EGVRYFVDHNTRTTTPKPRNGKSVTKGQPQIAYERGRWKLAFHRYLQSNALPSH	
			VKINVSQTLFEDSFOQIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFLLSHEVL	
			NPMYCLFEVAGKNNYCLQINPASTINPDHLSVFCFTGRFAMALFKGPKIDTGFSLPP	
			YKRLSKLTIKDLESIDTEFYNSLIWIRDNNEKGLSEMYFSDMEILLKGVTHDLK	
			LGSNLIIVTEENKDEYIGLMTWRPSRGVQEQTKAFIDGFNEVVPLOMQLVPEKELE	
			VMLCGMQEVDLADMQNRTVIRHTRNSKQIINFQVFKETDNEVRMLLQVFTGCKL	
			PLGFGAELMGNSGPKQKCIKRGVKOTWLPKSHRTCFNRLDLPYKYSEQLKEKLLFAIE	
			ETEGFGQE"	
ORIGIN				
Query Match	90.7%;	Score 566.2; DB 9; Length 3599;		
Best Local Similarity	97.3%;	Pred. No. 2e-144;		
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;			
Qy	27	CCCTTTAGTNCAGATCACTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	86	
Db	1924	CCTTTAGTTACAGATGACTTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	1865	
Qy	87	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	146	
Db	1864	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	1805	
Qy	147	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	206	
Db	1804	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	1745	
Qy	207	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	266	
Db	1744	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	1685	
Qy	267	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	326	
Db	1684	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	1625	
Qy	327	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	386	
Db	1624	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	1565	
Qy	387	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	446	
Db	1564	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	1505	
Qy	447	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	506	
Db	1504	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	1445	
Qy	507	CATGTGGTATGTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT	566	
Db	1444	CATGTGGTATGTCGAGCATATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT	1385	
Qy	567	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	624	
Db	1384	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	1328	

ORIGIN	/note="Region: HECT domain"		source	1. .3599
Query Match	90.7%;	Score 566.2; DB 9; Length 3479;	/organism="Homo sapiens"	
Best Local Similarity	97.3%;	Pred. No. 2e-144;	/mol_type="mRNA"	
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;		/db_xref="taxon:9606"	
			315_3083	
			/note="Tiul1; E3 ubiquitin ligase"	
			/codon_start=1	
			/product="TGIF-interacting ubiquitin ligase 1"	
			/protein_id="AAQ22764.1"	
			/db_xref="GI:33590476"	
			/translation="MATASPRSDTSNNHSRLQLOVTVSSAKLKRKNMFGTAIYEV	
			VIDGEITKTKASSSSNPWDQOLTVNTPQTLEQVWSHRTLKADALLKXKATIDLK	
			QALLIHNRLKRVKSLKSLNKGIAQBELTVVDGLVLIQSEINLTCSSPTTLEI	
			QENGDAHNGEPSEATLAVENGNGIDNHVPTSTLVQNSCCSVVSEVNDTPSPS	
			QVAAPKPTAPKPLASPADDTVNGESSFAPTDNASVTGPVSEVNSLNCSTST	
			TWDPVQIILTSSENNECI PSTSALESEARSILEPDTNSRSRFAEAKRQPDV	
			CMDVPQOSGNANTETLPSGMBQRKDPHGRTVVDHNTRTTWERPQPLPGWERRVD	
			DRRVYVVDHNTRTTWRPTWESVRNPEQWQSQRLQGMQFQRIYLSAMLA	
			ENDPGLPPEWEKRVDSIDRVYFNHNTKTQWEDPRTQGLQNEPFLPGWEIRYTR	
			EGVRYFVDHNTRTTTPKPRNGKSVTKGQPQIAYERGRWKLAFHRYLQSNALPSH	
			VKINVSQTLFEDSFOQIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFLLSHEVL	
			NPMYCLFEVAGKNNYCLQINPASTINPDHLSVFCFTGRFAMALFKGPKIDTGFSLPP	
			YKRLSKLTIKDLESIDTEFYNSLIWIRDNNEKGLSEMYFSDMEILLKGVTHDLK	
			LGSNLIIVTEENKDEYIGLMTWRPSRGVQEQTKAFIDGFNEVVPLOMQLVPEKELE	
			VMLCGMQEVDLADMQNRTVIRHTRNSKQIINFQVFKETDNEVRMLLQVFTGCKL	
			PLGFGAELMGNSGPKQKCIKRGVKOTWLPKSHRTCFNRLDLPYKYSEQLKEKLLFAIE	
			ETEGFGQE"	
ORIGIN				
Query Match	90.7%;	Score 566.2; DB 9; Length 3599;		
Best Local Similarity	97.3%;	Pred. No. 2e-144;		
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;			
Qy	27	CCCTTTAGTNCAGATCACTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	86	
Db	1924	CCTTTAGTTACAGATGACTTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	1865	
Qy	87	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	146	
Db	1864	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	1805	
Qy	147	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	206	
Db	1804	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	1745	
Qy	207	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	266	
Db	1744	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	1685	
Qy	267	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	326	
Db	1684	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	1625	
Qy	327	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	386	
Db	1624	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	1565	
Qy	387	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	446	
Db	1564	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	1505	
Qy	447	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	506	
Db	1504	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	1445	
Qy	507	CATGTGGTATGTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT	566	
Db	1444	CATGTGGTATGTCGAGCATATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT	1385	
Qy	567	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	624	
Db	1384	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	1328	

ORIGIN	/note="Region: HECT domain"		source	1. .3599
Query Match	90.7%;	Score 566.2; DB 9; Length 3479;	/organism="Homo sapiens"	
Best Local Similarity	97.3%;	Pred. No. 2e-144;	/mol_type="mRNA"	
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;		/db_xref="taxon:9606"	
			315_3083	
			/note="Tiul1; E3 ubiquitin ligase"	
			/codon_start=1	
			/product="TGIF-interacting ubiquitin ligase 1"	
			/protein_id="AAQ22764.1"	
			/db_xref="GI:33590476"	
			/translation="MATASPRSDTSNNHSRLQLOVTVSSAKLKRKNMFGTAIYEV	
			VIDGEITKTKASSSSNPWDQOLTVNTPQTLEQVWSHRTLKADALLKXKATIDLK	
			QALLIHNRLKRVKSLKSLNKGIAQBELTVVDGLVLIQSEINLTCSSPTTLEI	
			QENGDAHNGEPSEATLAVENGNGIDNHVPTSTLVQNSCCSVVSEVNDTPSPS	
			QVAAPKPTAPKPLASPADDTVNGESSFAPTDNASVTGPVSEVNSLNCSTST	
			TWDPVQIILTSSENNECI PSTSALESEARSILEPDTNSRSRFAEAKRQPDV	
			CMDVPQOSGNANTETLPSGMBQRKDPHGRTVVDHNTRTTWERPQPLPGWERRVD	
			DRRVYVVDHNTRTTWRPTWESVRNPEQWQSQRLQGMQFQRIYLSAMLA	
			ENDPGLPPEWEKRVDSIDRVYFNHNTKTQWEDPRTQGLQNEPFLPGWEIRYTR	
			EGVRYFVDHNTRTTTPKPRNGKSVTKGQPQIAYERGRWKLAFHRYLQSNALPSH	
			VKINVSQTLFEDSFOQIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFLLSHEVL	
			NPMYCLFEVAGKNNYCLQINPASTINPDHLSVFCFTGRFAMALFKGPKIDTGFSLPP	
			YKRLSKLTIKDLESIDTEFYNSLIWIRDNNEKGLSEMYFSDMEILLKGVTHDLK	
			LGSNLIIVTEENKDEYIGLMTWRPSRGVQEQTKAFIDGFNEVVPLOMQLVPEKELE	
			VMLCGMQEVDLADMQNRTVIRHTRNSKQIINFQVFKETDNEVRMLLQVFTGCKL	
			PLGFGAELMGNSGPKQKCIKRGVKOTWLPKSHRTCFNRLDLPYKYSEQLKEKLLFAIE	
			ETEGFGQE"	
ORIGIN				
Query Match	90.7%;	Score 566.2; DB 9; Length 3599;		
Best Local Similarity	97.3%;	Pred. No. 2e-144;		
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;			
Qy	27	CCCTTTAGTNCAGATCACTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	86	
Db	1924	CCTTTAGTTACAGATGACTTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	1865	
Qy	87	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	146	
Db	1864	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	1805	
Qy	147	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	206	
Db	1804	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	1745	
Qy	207	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	266	
Db	1744	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	1685	
Qy	267	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	326	
Db	1684	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	1625	
Qy	327	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	386	
Db	1624	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	1565	
Qy	387	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	446	
Db	1564	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	1505	
Qy	447	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	506	
Db	1504	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	1445	
Qy	507	CATGTGGTATGTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT	566	
Db	1444	CATGTGGTATGTCGAGCATATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT	1385	
Qy	567	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	624	
Db	1384	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	1328	

ORIGIN	/note="Region: HECT domain"		source	1. .3599
Query Match	90.7%;	Score 566.2; DB 9; Length 3479;	/organism="Homo sapiens"	
Best Local Similarity	97.3%;	Pred. No. 2e-144;	/mol_type="mRNA"	
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;		/db_xref="taxon:9606"	
			315_3083	
			/note="Tiul1; E3 ubiquitin ligase"	
			/codon_start=1	
			/product="TGIF-interacting ubiquitin ligase 1"	
			/protein_id="AAQ22764.1"	
			/db_xref="GI:33590476"	
			/translation="MATASPRSDTSNNHSRLQLOVTVSSAKLKRKNMFGTAIYEV	
			VIDGEITKTKASSSSNPWDQOLTVNTPQTLEQVWSHRTLKADALLKXKATIDLK	
			QALLIHNRLKRVKSLKSLNKGIAQBELTVVDGLVLIQSEINLTCSSPTTLEI	
			QENGDAHNGEPSEATLAVENGNGIDNHVPTSTLVQNSCCSVVSEVNDTPSPS	
			QVAAPKPTAPKPLASPADDTVNGESSFAPTDNASVTGPVSEVNSLNCSTST	
			TWDPVQIILTSSENNECI PSTSALESEARSILEPDTNSRSRFAEAKRQPDV	
			CMDVPQOSGNANTETLPSGMBQRKDPHGRTVVDHNTRTTWERPQPLPGWERRVD	
			DRRVYVVDHNTRTTWRPTWESVRNPEQWQSQRLQGMQFQRIYLSAMLA	
			ENDPGLPPEWEKRVDSIDRVYFNHNTKTQWEDPRTQGLQNEPFLPGWEIRYTR	
			EGVRYFVDHNTRTTTPKPRNGKSVTKGQPQIAYERGRWKLAFHRYLQSNALPSH	
			VKINVSQTLFEDSFOQIMALKPYDLRRLLYVIFRGEGLDYGGLAREWFLLSHEVL	
			NPMYCLFEVAGKNNYCLQINPASTINPDHLSVFCFTGRFAMALFKGPKIDTGFSLPP	
			YKRLSKLTIKDLESIDTEFYNSLIWIRDNNEKGLSEMYFSDMEILLKGVTHDLK	
			LGSNLIIVTEENKDEYIGLMTWRPSRGVQEQTKAFIDGFNEVVPLOMQLVPEKELE	
			VMLCGMQEVDLADMQNRTVIRHTRNSKQIINFQVFKETDNEVRMLLQVFTGCKL	
			PLGFGAELMGNSGPKQKCIKRGVKOTWLPKSHRTCFNRLDLPYKYSEQLKEKLLFAIE	
			ETEGFGQE"	
ORIGIN				
Query Match	90.7%;	Score 566.2; DB 9; Length 3599;		
Best Local Similarity	97.3%;	Pred. No. 2e-144;		
Matches 582; Conservative	0; Mismatches 15; Indels 1; Gaps 1;			
Qy	27	CCCTTTAGTNCAGATCACTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	86	
Db	1924	CCTTTAGTTACAGATGACTTCCCATTTGCGAGGATCTTTGAATCTGTGTGTTCTTGTTGTTA	1865	
Qy	87	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	146	
Db	1864	TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTGGCAGG	1805	
Qy	147	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	206	
Db	1804	GGTTCTTCATCTGTAAAGCCTTGAGTCTTGGATCTTCCACCTGGGTTGTTGTTGTTA	1745	
Qy	207	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	266	
Db	1744	TGATTCACAAAGTAAACCTCTGTGTGAATCCACTCTTTTCCAGCCTGTGGCAAA	1685	
Qy	267	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	326	
Db	1684	GGTCCATAAGGGTCATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA	1625	
Qy	327	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	386	
Db	1624	AACTGTTGATAGCTCCCTGCAATTTGGTTCGGCTGAGATGCCACCTGNTCAAAATNCGG	1565	
Qy	387	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	446	
Db	1564	ACAGATTCATGATAGCTAGCGCGCTGCGCAGCTTGTGTTCTGTGTTATGATCCACATATAA	1505	
Qy	447	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	506	
Db	1504	ACTCTCTACGATCATCAACTCTTCTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCTCC	1445	
Qy	507	CATGTGGTATGTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT	566	
Db	1444	CATGTGGTATGTCGAGCATATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT	1385	
Qy	567	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	624	
Db	1384	TGTTCCCAACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA	1328	

ORIGIN	/note="Region: HECT domain"				source
Query Match	90.7%;	Score 566.2;	DB 9;	Length 3479;	
Best Local Similarity	97.3%;	Pred. No. 2e-144;			
Matches 582;	Conservative 0;	Mismatches 15;	Indels 1;	Gaps 1;	
QY	27	CCCTTTAGTCAGATGACCTCCCATTCGGAGGATCTTTGAATGTTGTTGTTCTTGTTGA 86			
Db	1695	CCCTTTAGTTACAGATGACCTCCCATTCGGAGGATCTTTGAATGTTGTTGTTCTTGTTGA 1636			
QY	87	TGATCAACAAGTAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCACGCGCTTCGCGAG 146			
Db	1635	TGATCAACAAGTAGTACCTTTACACCTTCACGAGTATATCTAAATTTCCCGAGCTTCGCGAGG 1576			
QY	147	GGTTCTTCATCTGTGAAGCCTTGAGTTCTTGGATCTTTCCCACTGGGTGTTTGTGTTGA 206			
Db	1575	GGTTCTTCATCTGTGAAGCCTTGAGTTCTTGGATCTTTCCCACTGGGTGTTTGTGTTGA 1516			
QY	207	TGATTCACAAAGTAAACCTCTGCTGTGTAATCCACTCTTTTTTCCAGCGCTTCGTCGAGAA 266			
Db	1515	TGATTCACAAAGTAAACCTCTGCTGTGTAATCCACTCTTTTTTCCAGCGCTTCGTCGAGAA 1456			
QY	267	GCTCATTAAGGTCATTTTCTGCAGCTAACAATTGAAGCCGAATAGAGGTATCGTTGGTTA 326			
Db	1455	GCTCATTAAGGTCATTTTCTGCAGCTAACAATTGAAGCCGAATAGAGGTATCGTTGGTTA 1396			
QY	327	AACTCTTGTCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTCGG 386			
Db	1395	AACTCTTGTCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTCGG 1336			
QY	387	ACAGATTCATGTAGGCGCGTCGCCAGTTGTTGTTCTGGTGTATGATCCACATAATAA 446			
Ddb	1335	ACAGATTCATGTAGGCGCGTCGCCAGTTGTTGTTCTGGTGTATGATCCACATAATAA 1276			
QY	447	ACTCCTCTACAGATCATCAACTCTTTTCCCACTCGAGGTAAAGTTGTTGTTCTCTCC 506			
Ddb	1275	ACTCCTCTACAGATCATCAACTCTTTTCCCACTCGAGGTAAAGTTGTTGTTCTCTCC 1216			
QY	507	CATGTGGTAGTTTCAGAGCATTAATGANCACATAATACCTCTACCATGAGGATCTTTTCTT 566			
Ddb	1215	CATGTGGTAGTTTCAGAGCATTAATGATCCACATAATAGTTTCTACATGAGGATCTTTTCTT 1156			
QY	567	TGTTCCCACTGATGCGCAANGNTTCTGTGTTGGCATTTCCAGACTGTCGCGNACA 624			
Ddb	1155	TGTTCCCA-CCCTGATGCGCAAGGTTTCTGTGTTGGCATTTCCAGACTGTCGCGGTACA 1099			
RESULT 7					
AY345857/c					
LOCUS					
DEFINITION	Homo sapiens TGIF-interacting ubiquitin ligase 1 mRNA, complete cds.				
ACCESSION	AY345857				
VERSION	AY345857.1				
KEYWORDS	GI:33590475				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 3599)					
REFERENCE	Seo, S.R., Lallemand, F., Ferrand, N., Pessah, M., L'Hoste, S., Camonis, J. and Atfi, A.				
AUTHORS	The novel E3 ubiquitin ligase Tiul1 associates with TGIF to target Smad2 for ubiquitin-mediated degradation				
TITLE	Unpublished				
JOURNAL	2 (bases 1 to 3599)				
REFERENCE	Seo, S.R., Lallemand, F., Ferrand, N., Pessah, M., L'Hoste, S., Camonis, J. and Atfi, A.				
AUTHORS	Direct Submission				
TITLE	Submitted (18-JUL-2003) INSERM U482, Hôpital Saint-Antoine, 184 Rue du Faubourg St-Antoine, Paris 75571, France				
JOURNAL	Location/Qualifiers				
AUTHORS	BEATURES				

/gene="WWP1"
/note="synonyms: AIP5, hSDRP1, DKFZp434D2111, Tiul1"
/db_xref="LocusID:11059"
/db_xref="MIM:602307"
285. .3053
/gene="WWP1"
/codon_start=1
/product="WW domain containing E3 ubiquitin protein ligase
1"
/protein_id="AAH36065.1"
/db_xref="GI:23271281"
/db_xref="LocusID:11059"
/db_xref="MIM:602307"
/translation="MATA SPRSDTSNNHSGRLQVTVVSASAKLRKKNFNGTALTYEV
VDPGETITAKSSSSNPKWDEQLTVNVTPTQTFEFQVMSHRTLKADALLGKATIDLK
OALLHNKLERVKEOLKLSLENGKIAOTGVLTVDGLVIGQENITNCSSPTTIEI
QENGDAHLNGEFSARTTARLAVETGNGIDNVHTSTLVNSCCSVVANGDNPSPSP
QVAPRPNTAPAPLSEADDDTVNGESSFAPTDNASVTGTSPVSEENALSPNCTST
TVEDPPVOEILTSSNNECIPSPSAELESARSILPDTNSRSSSAFAAKSRORDG
CMDFVROQSGNANTETLPSGWEQRKDPHGRTTYVDHNTTTRTWERQPLPGWHRVD
DRRVVYVDHNTTTRTWERQPTMESVRNFQWQSRNOLQAOQFQNRQYIYSASMLAA
ENDPYGVLDHNTTTRTWERQPTMESVRNFQWQSRNOLQAOQFQNRQYIYSASMLAA
EGVRYVDHNTTTRTKDPRNGKSVVTKGGPQIAYERGRWKLAFRLYLCQSALPESH
VKINVSQTLFEDSPFOQIMALKPVDLRRLLVIFRGEGLDYGLAREWFLLSHEVLI
NPMYCLFAGAKNNYCIQINPASTINPDHLSYECFGRFETAMALFHGKIDTGFSLPF
YKRLSKLTIKDLESIDTFYNSLIWIRDNNEIEECGLEMYFSDVMEILLKGVSHDLK
LGSSNLIIVTEENKDEYIGLMTFRFSGVQEQTKAFIDGNEVVPVQWLOLQYFDEKLE
VMLCGMOVDLADQNRNTVYRHTNSKQIIFWQFVKETDNEVRMLLQFVGTGRL
PLGGFAELMGSNGPKCEIKVGKDTWLPRSHTCFNRLDLPYKYSVQOLKEKLLFAIE
ETEGFGQE"

CDS

ORIGIN

Query Match	Best Local Similarity	Score	566.2;	DB 9;	Length	3855;
Matches	Conservative	0;	Mismatches	15;	Indels	1;
QY	27	CCCTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTGTTGTTA	86			
DB	1894	CTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTGTTA	1835			
QY	87	TGATCAACAAAGTACCTTACACCTTACAGATATATCTAATTTCCAGCCTTCTGGCAGG	146			
DB	1834	TGATCAACAAAGTACCTTACACCTTACAGATATATCTAATTTCCAGCCTTCTGGCAGG	1775			
QY	147	GGTTCTTCTTCTGTAAGCCTTGAATCTTCCCATTCGAGGATCTTTGTTGTTGTTA	206			
DB	1774	GGTTCTTCTTCTGTAAGCCTTGAATCTTCCCATTCGAGGATCTTTGTTGTTGTTA	1715			
QY	207	TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCACGCTGGTGGCAA	266			
DB	1714	TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTCACGCTGGTGGCAA	1655			
QY	267	GGTCCATAAGGGTCATTTTCGAGCTAACATTAAGCCGATAGAGGTATCGTTGGTTA	326			
DB	1654	GGTCCATAAGGGTCATTTTCGAGCTAACATTAAGCCGATAGAGGTATCGTTGGTTA	1595			
QY	327	AACGTGTCATAGCTCCCTGCAATTCGTTCCGCTGAGATTGCCACTGNTCAAAATTCGG	386			
DB	1594	AACGTGTCATAGCTCCCTGCAATTCGTTCCGCTGAGATTGCCACTGNTCAAAATTCGG	1535			
QY	387	ACAGATTCCATGGTAGGCGCTGCCACGTTGTTGTTGTTGTTGTTGTTGTTGTTA	446			
DB	1534	ACAGATTCCATGGTAGGCGCTGCCACGTTGTTGTTGTTGTTGTTGTTGTTGTTA	1475			
QY	447	ACTCTCTACGATCATCACTCTTCTTCCCACTCGAGGTAAGAGTTGTTGTTCTCTCC	506			
DB	1474	ACTCTCTACGATCATCACTCTTCTTCCCACTCGAGGTAAGAGTTGTTGTTCTCTCC	1415			
QY	507	CATGTGTTAGTTTCAGCATTTATGANCACATAATAACGCTCTACATGAGGATCTTTCTT	566			
DB	1414	CATGTGTTAGTTTCAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTCTT	1355			
QY	567	TGTTCCCAACCCCTGATGGCAANGTCTCTGTGTTGGCATTTCCAGACTGTGCGGNACA	624			

RESULT 8

BC036065/c
LOCUS
DEFINITION
mRNA (CDNA clone MGC:33686 IMAGE:5296005), complete cds.

ACCESSION

BC036065

VERSION

BC036065.1 GI:23271280

KEYWORDS

MGC

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 3855)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, P.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
Ackerman, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. M.A. A full-length
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3855)
Strausberg, R.
Direct Submission
Submitted (31-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki
Toehiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILLNL at: <http://image.llnl.gov>
Series: IRAC plate: 48 Row: K Column: 20
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19718739.

FEATURES

source

1. .3855
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:33686 IMAGE:5296005"
/tissue_type="Testis"
/clone_lib="NIH MGC_97"
/lab_host="DH10B"
/notes="Vector: pBluescript"

gene

1. .3855

Db 1354 TGTTCCTCA-CCTGATGCGAAGGTTCTGTGTGGCATTCCTCCAGACTGCTGCCGTACA 1298
|||||
RESULT 9
AX053142/c
LOCUS AX053142.1 3494 bp DNA linear PAT 13-JAN-2001
DEFINITION Sequence 4 from Patent WO0073329.
ACCESSION AX053142
VERSION AX053142.1 GI:12227506
KEYWORDS
SOURCE Drosophila sp.
ORGANISM Drosophila sp.
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscophila;
Ephydroidea; Drosophilidae; Drosophila.
1
REFERENCE
AUTHORS Baron, M.
TITLE Uses of notch related genes
JOURNAL Patent: WO 0073329-A 4 07-DEC-2000;
THE VICTORIA UNIVERSITY OF MANCHESTER (GB)
FEATURES
source Location/Qualifiers
1..3494
/organism="Drosophila sp."
/mol_type="unassigned DNA"
/db_xref="taxon:7242"
ORIGIN
Query Match 88.58; Score 552.2; DB 6; Length 3494;
Best Local Similarity 97.08; Pred. No. 1.4e-140;
Matches 580; Conservative 0; Mismatches 15; Indels 3; Gaps 2;
QY 27 CCCTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTTA 86
|||
Db 1710 CCTTGTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCT -TGTTA 1653
|||
QY 87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 146
|||
Db 1652 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 1593
|||
QY 147 GGTCTCTCATCTGTAGAGCTTGGATCTTGGATCTTCCACATGGGTTGTTTGTGTTA 206
|||
Db 1592 GGTCTCTCATCTGTAGAGCTTGGATCTTGGATCTTCCACATGGGTTGTTTGTGTTA 1533
|||
QY 207 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTCCAGCCTGTGGCAAA 266
|||
Db 1532 TGATTCACAAAGTAAACCTCTGTGTTGAATCCACTCTTTTCCAGCCTGTGGCAAA 1473
|||
QY 267 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCCGAATAGAGGTATCGTTGGTTA 326
|||
Db 1472 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCCGAATAGAGGTATCGTTGGTTA 1413
|||
QY 327 AACTGTTGATAGTCCCTGCAATTTGGTTCGGCTGAGATGGCCATGNTCAAAATTCGG 386
|||
Db 1412 AACTGTTGATAGTCCCTGCAATTTGGTTCGGCTGAGATGGCCATGNTCAAAATTCGG 1353
|||
QY 387 ACAGATTCATGTTAGGCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATATAA 446
|||
Db 1352 ACAGATTCATGTTAGGCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATATAA 1293
|||
QY 447 ACTCCTCTACGATCATCAACTCTTTTCCCAACCTCGAGGTAAAGTTGTTGTTCTCC 506
|||
Db 1292 ACTCCTCTACGATCATCAACTCTTTTCCCAACCTCGAGGTAAAGTTGTTGTTCTCC 1233
|||
QY 507 CATGTGTAGTTCGAGCATTTATGANCACATATACCTCTACCATGAGGATCTTTCTT 566
|||
Db 1232 CATGTGTAGTTCGAGCATTTATGATCCACATATAGTTCTTACCATGAGGATCTTTCTT 1173
|||
QY 567 TGTTCCTCCACCTCATGTCGAAGGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNACA 624
|||
Db 1172 TGTTCCTCA-CCTGATGCGAAGGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGTACA 1116
|||||

RESULT 10
AC016962
LOCUS AC016962
DEFINITION Homo sapiens 3 BAC RP11-569H14 (Roswell Park Cancer Institute Human BAC Library) complete sequence.
ACCESSION AC016962.38 GI:21206149
VERSION AC016962.38
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 49548)
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Alisbrooks, S.L., Amaratunge, H.C., Are, J.R., Ayele, M., Banks, T.,
Barbata, J., Benton, J., Bimaga, K., Blankenburg, K., Bonnin, D.,
Bouck, J., Bowie, S., Brieva, M., Brown, E., Brown, M., Bryant, N.P.,
Buhay, C., Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C.,
Carron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D.,
Chen, G., Chen, R., Chen, Z., Chowdhry, I., Christopoulos, C.,
Cleveland, C.D., Cox, C., Coyie, M.D., Dathorne, S.R., David, R.,
Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A.,
Delaney, K.R., Delgado, O., Denn, A.L., Ding, Y., Dinh, H.H.,
Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J.,
Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Escotto, M.,
Falls, T., Ferraguto, D., Flagg, N., Ford, J., Foster, P., Frantz, P.,
Gabisi, A., Gao, J., Garcia, A., Garner, T., Garza, N., Gill, R.,
Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S., Hamilton, K.,
Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A., He, X.,
Hernandez, J., Hernandez, O., Hodgson, A., Hogues, M., Holloway, C.,
Hollins, B., Honsi, F., Howard, S., Huber, J., Hui, J., Hume, J.,
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,
Loulseghe, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,
Martinez, E., Massey, E., Mawhiney, E., McLeod, M.P., Meador, M.,
Mei, G., Metzker, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,
Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,
Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,
Nickerson, E., Nwokenko, S., Ogih, M., Okunolu, G., Oragunye, N.,
Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,
Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y., Rives, M.,
Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G., Scherer, S.,
Scott, G., Shen, H., Shoshitari, N., Sisson, I., Sodergren, E.,
Sonaite, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,
Taylor, T., Telford, B., Thomas, N., Thomas, S., Tansey, J., Taylok, C.,
Vera, V., Villalon, D., Vinson, R., Wang, Q., Wang, S., Ward-Moore, S.,
Warren, R., Washington, C., Watlington, S., Williams, G.,
Williamson, A., Wlezyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,
Wu, Y.F., Zhou, J., Zorrilla, S., Naylor, S.L., Weinstein, G. and
Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (09-DEC-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (09-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
4 (bases 1 to 49548)
Worley, K.C.
Direct Submission
Submitted (25-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One


```
QY 207 TGATTCAAAAGTAACCCCTGTCGTGTAATCCACTCTTTTCCAGCCCTGGTGGCAAA 266
Db 45182 CGATTCAAAAGTAACCCCTGTCGTGTAATCCACTCTTTTCCAGCCCTGGTGGCAAA 45241
QY 267 GGTCCATAAGGGTCATTTCTGCGAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGGTTA 326
Db 45242 GGTCCATAAGGGTCATTTCTGCGGCTAAACATTGAAGCTGAATAGAGGTATCATTTGGTTA 45301
QY 327 AACTGTTGCTAGTAGCTCCCTGCAATTTGTTTCGGCTGAGATTGCCACTGNTCAAAATTCGCG 386
Db 45302 AACTGTTGCTAGTAGCTCCCTGAGGTTGTTCTGCTGAGATTGCCACTGTTCAAAATTTCTG 45361
QY 387 ACAGATTCCATGTTAGCCCGCTGCCAGCTGTTGTTGTTCTGTTGTTATGATCCACATATAA 446
Db 45362 ACCAATTCATGTTAGCCCGCTGCCATGTTGTTCTTCTGTTGTTATGATCCACATATAA 45421
QY 447 ACTCTCTAGCATCATCACTCTCTCTTCCCACTCGAGGTAAAGGTTGTTGTTCTCTCC 506
Db 45422 GCTCTTCCAGCATCATCACTCTCTTCTTCCCACTCGAGGTAAAGGTTGTTGTTCTCTCC 45481
QY 507 CATGTGTTAGTTTGCAGCATTTATGANCACATATAATACCTCTTACCATGAGGATCTTTTCTT 566
Db 45482 CGTGTGTTAGTTTGCAGTTATGATCCACATATAGTTTCTACCATGAGGATTTTCTT 45541
QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNAC 623
Db 45542 TGTTCACCA-CCCTGATGGCAAGGTTTCTGTGTTGGCATTCCTCCAGACTCCTGCTATAC 45597

RESULT 11
AF412122/c
LOCUS AF412122 1794 bp mRNA linear VRT 12-MAR-2002
DEFINITION Gallus gallus late domain-interacting protein 2 mRNA, partial
sequence.
ACCESSION AF412122.1 GI:19401482
VERSION AF412122.1
KEYWORDS Gallus gallus (chicken)
SOURCE Gallus gallus
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 1794)
Kikonyogo, A., Bouamr, F., Vana, M.L., Xiang, Y., Aliyar, A., Carter, C.
and Leis, J.
Proteins related to the Nedd4 family of ubiquitin protein ligases
interact with the L domain of Rous sarcoma virus and are required
for gag budding from cells
Proc. Natl. Acad. Sci. U.S.A. 98 (20), 11199-11204 (2001)
MEDLINE 21457266
PUBMED 11562473
REFERENCE 2 (bases 1 to 1794)
Kikonyogo, A., Bouamr, F., Vana, M.L., Xiang, Y., Aliyar, A., Carter, C.
and Leis, J.
Direct Submission
Submitted (22-AUG-2001) Microbiology and Immunology, Northwestern
University, 303 E. Chicago, Chicago, IL 60611, USA
FEATURES
source
location/Qualifiers
1..1794
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
100..>1794
/notes="similar to E3 ubiquitin ligase; late
domain-interacting protein 2; LDI-2; coding region not yet
determined"
misc_feature 1162..1251
/notes="potential WW 1"
misc_feature 1255..1347
/notes="potential WW 2"
misc_feature 1483..1569
/notes="potential WW 3"
misc_feature 1597..1692
```

ORIGIN

Query Match 69.1%; Score 431.2; DB 5; Length 1794;
Best Local Similarity 81.8%; Pred. No. 2.5e-107;
Matches 504; Conservative 0; Mismatches 111; Indels 1; Gaps 1;

/note="potential WW 4"

```
QY 9 TTNTAAGCAATTTGTGGACCCCTTTAGTNCAGATGACTTCCCAATTCGAGGATCTTTCAAT 68
Db 1736 TCATAGGCAATCTGTGGCCCTTTTATTATTACAGAAGACTTCCAGTACGAGGATCATTTGAAG 1677
QY 69 GTTGTGTTCTTGTGTTATGATCAACAAAGTACTTACACCTTACACGAGTATATCTAATT 128
Db 1676 GTGGTAGTCTTGTATATGATCAACAAAGTATCGTACACCTTCTCTAGTGTATCTGATT 1617
QY 129 TCCAGAGCTTCTGCGAGGGTCTTCAATCTGTAAAGCTTTGAGTCTTGTGATCTTCCAC 188
Db 1616 TCCAGAGCTTCTGCGAGGGTCTCTCAATTTGTAAACCTTTGGTTCGAGGGTCTTCCAT 1557
QY 189 TGGGTTGTTTGTGTTATGATTACAAAGTAAACCTCTGTCTGTGAATCCACTCTTTTT 248
Db 1556 TGTGTTGTCTTGTGTTATGATTACAAATATACACCTATCATTTGAATCCACTCTCTT 1497
QY 249 TCCAGAGCTGTTGCGAAAGTCCATTAAGGTCATTTTCTGAGCTAATTAAGCCGAA 308
Db 1496 TCCAGAGCTGTTGCGAAAGCCCAAGTGGATCGTTCTCTGCTGACACATCGAAGCCGAA 1437
QY 309 TAGAGGTATCTGTTGGTTAAACTGTTCATAGCTCCCTGCAATTTGTTCCGCTGAGATTGC 368
Db 1436 TAGAGGTATCTGTTGGTTAAATTTGTTCATAGCTCCCTGCAATTTGATTAACCTGAGATGC 1377
QY 369 CACTGNTCAAAATTCGCGAGGATTCATGTTAGGCGCTGCCACGTTGTTGTTCTGGTG 428
Db 1376 CACTGCTCAAAAGTCTCTGACTGATTCCATTGTTGGTGGCTGCCATGTTGTTGTTCTG 1317
QY 429 TTATGATCCACATATAATAACTCTCTACGATCATCAACTCTTCTTCCCAACCTCGAGGT 488
Db 1316 TTATGTTCCACATAGTAAACTCTTCCAGCATCATCAACTCTTCTTCCAGCCAGGAGGC 1257
QY 489 AAAGGTTGTTGTTCTCTCCCATGTTGTTGAGTATTCAGAGCATTAAGNCCACATATAAGCTTA 548
Db 1256 AAGGGCTGTTGTTCTTCCCACTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1197
QY 549 CCATGAGGATCTTTCTTTGTTTCCCACTGATGCGCAANGTTCGTGTGTCATTTCCC 608
Db 1196 CCATGAGGATCTTTCTTTGTTTCCCA-CCCTGTTGTTAAGGTTCTGTACCAGCATTAAC 1138
QY 609 AGACTGCTGCCGNACA 624
Db 1137 AGTCTGCTGCTTACA 1122
```

RESULT 12

BC076842/c
LOCUS BC076842 3501 bp mRNA linear VRT 19-JUL-2004
DEFINITION Xenopus laevis cDNA clone IMAGE:6859998, partial cds.

ACCESSION BC076842

VERSION BC076842.1 GI:50368904

KEYWORDS

SOURCE Xenopus laevis (African clawed frog)

ORGANISM

Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.

REFERENCE

AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W. and Richardson, P.

TITLE

Genetic and genomic tools for Xenopus research: The NIH Xenopus initiative

JOURNAL

PUBMED Dev. Dyn. 225 (4), 384-391 (2002)

REFERENCE

2 (bases 1 to 3501)

AUTHORS

Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,

Klausner, R.D., Collins, F.S., Wagner, L., Shennan, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, T.E., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Schetz, T.M., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzyzinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
3 (bases 1 to 3501)
Klein, S. and Strausberg, R.
Direct Submission
Submitted (06-JUL-2004) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
NIH-MGC Project
Contact: XGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield, Andy Chan, Steve S. Chand, William Chow, Allison Cloutier, Ruth Featherstone, Malachi Griffith, Obi Griffith, Ran Guin, Nancy Liao, Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin, Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lilisa Prahbu, Parvaneh Saeedi, JR Santos, Angélique Schnerch, Ursula Skalska, Duane Smallos, Jeff Stott, Miranda Tsai, George Yang, Jacques Schein, Asim Siddiqui, Rob Holt, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 158 Row: 0 Column: 10
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source
1. .3501
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:685998"
/issue_type="Oocytes"
/clone_lib="NICHDR_XGC_001"
/lab_host="DH108"
/note="Vector: pCMV-SPORT6.1"

ORIGIN

Query Match 58.9%; Score 367.8; DB 5; Length 3501;
Best Local Similarity 79.7%; Pred. No. 5.9e-90;
Matches 432; Conservative 0; Mismatches 110; Indels 0; Gaps 0;
QY 37 CAGATGACTTCCCATTCGCGAGGATCTTTGAATCTGTCTTCTTGTATGATCAACAA 96
Db 1837 CTGCTGACTTCCCGTTGCTGGGTCGCTGAAGTAGTTGCTTTGATTTGATCTACAA 1778
QY 97 AGTACCTTACACCTTCACAGATATATCTAATTTCCCGAGCTTCTGCGAGGGTTCTTCAT 156

Db 1777 AGTACTTTACTCTCTCCCTTGTGTAAGGATTTCCCGAGCCCTCTGGCAATGATCTCAT 1718
QY 157 TCTGTAAGCCTTGAGTCTTCTGATCTTCCCATCTGGTGTGTTTGTGTTATGATCAAA 216
Db 1717 TCTGTAAGCCTTGAGTCTTCTGATCTTCCCATCTGGTGTGTTTGTGTTATGATCAAA 1658
QY 217 AGTAAACCCCTGCTGTGTAATCACTCTTTTTTCCAGCTCTGGTGCAAGGTCCATAAG 276
Db 1657 AATAGACCCGATCAGTGGATCCATCTCTCTCCCATCTGGCGGAAGTGTCCCAAG 1598
QY 277 GGTCAATTTTCTCAGCTAACTTGAAGCGAATAGAGGTATCGTGGTTAAACTGTTGCA 336
Db 1597 GGTCAATTTTCTCAGCGGATCACTTGAATGCGAATAGAGGTATCGTGGTTAAACTGTTGCA 1538
QY 337 TAGCTCCCTGCAATTTGGTTCCTGATGATGCGACATGNTCAAAATTCGACAGATTCGA 396
Db 1537 TAGCTCCCTGCAATTTGGTTCCTGATGATGCGACATGNTCAAAATTCGACAGATTCGA 1478
QY 397 TGGTAGGCGCTGCCACAGTTGTTGTTCTGGTGTATGATCCACATAATAAACTCTCTAC 456
Db 1477 TAGTAGGCTCTGCCATGTTGTTGTTCTGTTGTTATGTTCAACATAATACACTCTCCAC 1418
QY 457 GATCATCAACTCTTCTTTCCCAACCTGAGGTAAAGGTTGTGTTCTCTCCCATGTTGTTAG 516
Db 1417 GATCATCGACCTCTCTCTCCCAACCTGAGGTAAAGGTTGTGTTCTCTCCCATGTTGTTAG 1358
QY 517 TTCGAGCATATGANCACACATTAATGCTCTACCATGAGATCTTTCTTTGTTCCCAAC 576
Db 1357 TCCTTGTGTTATGATCAACGTAGTAGTTCTACCATGGGATCTTTCTCTGTTTCCCAAC 1298
QY 577 CC 578
Db 1297 CC 1296

RESULT 13

AC101726 130034 bp DNA linear HTG 11-JUN-2003
LOCUS Mus musculus clone RP23-321024, WORKING DRAFT SEQUENCE, 10
DEFINITION unordered pieces.
AC101726
AC101726.3 GI:31581761
HTG; HTGS PHASE1; HTGS DRAFT; HTGS_FULLTOP.
VERSION Mus musculus (house mouse)
SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 130034)

REFERENCE

Birren, B., Nusbaum, C. and Lander, E.

AUTHORS

Mus musculus, clone RP23-321024

TITLE

Unpublished

JOURNAL

2 (bases 1 to 130034)

REFERENCE

Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, I., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., DeArelano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Levine, R., Liu, G., Lamazares, R., Lander, T., Lehoczy, J., Levine, R., Liu, G., MacLean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, I., Mihova, T., Mienga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strause, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H.,

Db 76124 CTTTCCCAACTGGAGCA 76142
|||||

```

RESULT 14
AC103927
LOCUS
DEFINITION
AC103927
ACCESSION
AC103927.4
VERSION
GTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLPOP.
KEYWORDS
MUS musculus (house mouse)
SOURCE
- ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,
Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S.,
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kanat,A., Karatas,A., Kellis,C., Lakocque,K.,
Lanazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,
McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrim,J.,
Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D.,
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,
Roman,J., Roberti,M., Roy,A., Santos,R., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
Strauss,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 135660)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,M., Atachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
Boguslavskiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Corum,B., Dearellano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kellis,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,X., Liu,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramaeamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (09-MAR-2004) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 28, 2004 this sequence version replaced gi:31455749.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L17179
Center clone name: 44 A.16
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 131840 bases at least Q40
Consensus quality: 132850 bases at least Q30
Consensus quality: 133337 bases at least Q20
Insert size: 127000; agarose-fp
Insert size: 134128; sum-of-contigs
Quality coverage: 12.1 in Q20 bases; agarose-fp
Quality coverage: 11.5 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
13935: contig of 13935 bp in length
13936 14035: gap of 100 bp
14036 15317: contig of 1282 bp in length
15318 15417: gap of 100 bp
15418 17222: contig of 1805 bp in length
17223 17322: gap of 100 bp
17323 22104: contig of 4782 bp in length
22105 22204: gap of 100 bp
22205 31776: contig of 9572 bp in length
31777 39882: contig of 8006 bp in length
39883 39982: gap of 100 bp
39983 65137: contig of 25154 bp in length
65137 65237: gap of 100 bp
65237 75051: contig of 9815 bp in length
75052 75151: gap of 100 bp
75152 92728: contig of 17577 bp in length
92729 92828: gap of 100 bp
92829 115139: contig of 22311 bp in length
115140 115240: gap of 100 bp
115240 116031: contig of 792 bp in length
116032 116131: gap of 100 bp
116132 135660: contig of 19529 bp in length.

FEATURES

	Location/Qualifiers	source
	1..135660	
	/organism="Mus musculus"	
	/mol_type="genomic DNA"	
	/db_xref="taxon:10090"	
	/clone="RP23-44A16"	
	/clone_lib="RPCI-23 Female Mouse BAC"	
	1..13935	
misc_feature	/note="assembly_fragment"	
	clone_end:SP6	
	vector side:left	
	14036..15317	
misc_feature	/note="assembly_fragment"	
	15418..17222	
misc_feature	/note="assembly_fragment"	
	17323..22104	
misc_feature	/note="assembly_fragment"	
	22205..31776	
misc_feature	/note="assembly_fragment"	
	31877..39882	
misc_feature	/note="assembly_fragment"	

* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

```

1 692: contig of 692 bp in length
693 792: gap of 100 bp
793 1433: contig of 641 bp in length
1434 1533: gap of 100 bp
1534 2667: contig of 1134 bp in length
2668 2767: gap of 100 bp
2768 3958: contig of 1191 bp in length
3959 4059: gap of 100 bp
4060 4761: contig of 703 bp in length
4762 4862: gap of 100 bp
4863 6373: contig of 1512 bp in length
6374 7488: contig of 1015 bp in length
7489 8332: gap of 100 bp
8333 8432: gap of 100 bp
8433 10101: contig of 1669 bp in length
10102 10201: gap of 100 bp
10202 11524: contig of 1323 bp in length
11525 11625: gap of 100 bp
11626 13293: contig of 1669 bp in length
13294 13393: gap of 100 bp
13394 14364: contig of 971 bp in length
14365 16117: contig of 1653 bp in length
16118 16217: gap of 100 bp
16218 17410: contig of 1193 bp in length
17411 17510: gap of 100 bp
17511 19574: contig of 2064 bp in length
19575 19674: gap of 100 bp
19675 21717: contig of 2042 bp in length
21718 21816: gap of 100 bp
21817 24076: contig of 2260 bp in length
24077 24176: gap of 100 bp
24177 27832: contig of 3656 bp in length
27833 27932: gap of 100 bp
27933 48374: contig of 20442 bp in length
48375 48474: gap of 100 bp
48475 63552: contig of 15078 bp in length
63553 63653: gap of 100 bp
63654 101162: contig of 37510 bp in length
101163 101262: gap of 100 bp
101263 189541: contig of 88279 bp in length.

FEATURES             Location/Qualifiers
     source            1..189541
                        /organism="Mus musculus"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:10090"
                        /clone="RP24-136P4"
                        /clone_lib="RPCI-24 Male Mouse BAC"
     misc_feature      1..692
                        /note="assembly_fragment"
     misc_feature      793..1433
                        /note="assembly_fragment"
     misc_feature      1534..2667
                        /note="assembly_fragment"
     misc_feature      2768..3958
                        /note="assembly_fragment"
     misc_feature      4059..4761
                        /note="assembly_fragment"
     misc_feature      4862..6373
                        /note="assembly_fragment"
     misc_feature      6474..7488
                        /note="assembly_fragment"
     misc_feature      7589..8332
                        /note="assembly_fragment"
     misc_feature      8433..10101
                        /note="assembly_fragment"
     misc_feature      10202..11524
                        /note="assembly_fragment"

```

```

misc_feature          11625..13293
                        /note="assembly_fragment"
misc_feature          13394..14364
                        /note="assembly_fragment"
misc_feature          14465..16117
                        /note="assembly_fragment"
misc_feature          16218..17410
                        /note="assembly_fragment"
misc_feature          17511..19574
                        /note="assembly_fragment"
misc_feature          19675..21716
                        /note="assembly_fragment"
misc_feature          21817..24076
                        /note="assembly_fragment"
misc_feature          24177..27832
                        /note="assembly_fragment"
misc_feature          27933..48374
                        /note="assembly_fragment"
misc_feature          48475..63552
                        /note="assembly_fragment"
misc_feature          63653..101162
                        /note="assembly_fragment"
misc_feature          101263..189541
                        /note="assembly_fragment"

ORIGIN

```

```

Query Match          57.0%; Score 355.4; DB 2; Length 189541;
Best Local Similarity 90.9%; Pred. No. 8.4e-87;
Matches 399; Conservative 0; Mismatches 38; Indels 2; Gaps 2;

QY 51 TTGCGAGGATCTTGAATGTTGTTCTTCTGCGAGGGGTTCTTCTTCTGTTATGATCAACAAGTACCTTACACCT 110
Db 29461 TTCCAGGATCTTGAATGTTGTTCTTCTGCGAGGGGTTCTTCTTCTGTTATGATCAACAAGTACCTTACACCT 29403

QY 111 TCACGAGTATPCTAATTTCCAGCCTTCTGCGAGGGGTTCTTCTTCTGTTATGATCAACAAGTAAACCTTGTCT 170
Db 29402 TCCCGAGTATACCTAATTTCCAGCCTTCTGCGAGGGGTTCTTCTTCTGTTATGATCAACAAGTAAACCTTGTCT 29343

QY 171 GTTCTTGGATCTTCCCACTGGGTTGTTTCTGTTATGATCAACAAGTAAACCTTGTCT 230
Db 29342 GTTCTTGGATCTTCCCACTGGGTTGTTTCTGTTATGATCAACAAGTAAACCTTGTCT 29283

QY 231 GTTGAATCCACTCTTTTTCAGCCTGTTGCGAAAGTCCATTAAGGTCATTTTCTGCA 290
Db 29282 GTTGAATCCACTCTTTTTCAGCCTGTTGCGAAAGTCCATTAAGGTCATTTTCTGCA 29223

QY 291 GCTAACATTGAAGCGGAATAGAGTATCGTTGGTTAAACTGTTGATAGCTCCCTGCAAT 350
Db 29222 GCTAACATTGAAGCGGAATAGAGTATCGTTGGTTAAACTGTTGATAGCTCCCTGCAAT 29164

QY 351 TGGTTCCGCTGAGATTGCGACTGNTCAAAATTCGGACAGATTCATGTTAGGCGGTGC 410
Db 29163 TGGTTCCGCTGAGATTGCGACTGNTCAAAATTCGGACAGATTCATGTTAGGCGGTGC 29104

QY 411 CACGTTGTTGTTGTTGTTATGATCAACAATAAATCTCTTACGATCATCAACTCTT 470
Db 29103 CAGGTTCTTCTGTTGTTGTTATGATCAACAATAAATCTCTTACGATCATCAACTCTT 29044

QY 471 CTTTCCCAACTCGAGGTA 489
Db 29043 CTTTCCCAACTCGAGGTA 29025

```

Search completed: June 30, 2005, 05:30:24
 Job time : 3093 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 03:18:24 ; Search time 471 Seconds
(without alignments)
7842.715 Million cell updates/sec

Title: US-10-618-408-3

Perfect score: 624
Sequence: 1 ccgcgcgtntaagcattt.....tccagactgctgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10_0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2002bs:*

8: Geneseqn2003as:*

9: Geneseqn2003bs:*

10: Geneseqn2003cs:*

11: Geneseqn2003ds:*

12: Geneseqn2004as:*

13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	616	98.7	624	13	ADS00022 Human WWP
2	566.2	90.7	2052	2	AAT95697 Novel hum
3	566.2	90.7	2052	4	AAD10504 Human ubi
4	566.2	90.7	2052	6	ABZ34944 Human gen
5	566.2	90.7	2052	8	AAD50455 Nedd-4-1i
6	566.2	90.7	2052	10	ADB49239 Novel hum
7	566.2	90.7	2731	10	ADC26288 Human NOV
8	566.2	90.7	2793	13	ADS00020 Human WWP
9	566.2	90.7	2827	4	AAC90422 Human clo
10	566.2	90.7	3382	6	ABA93745 Human sig
11	563.6	90.3	943	13	ACN38827 Tumour-as
12	552.2	88.5	3495	4	AAC90423 Human hom
13	456.8	73.2	2305	12	ADP84403 Human bre
14	246.6	39.5	2217	5	AAS80664 DNA encod
15	246.6	39.5	2377	5	ABA82679 Atrophin-
16	246.6	39.5	2377	8	AAD50454 Atrophin-
17	246.6	39.5	2377	8	ACC45999 Human atr
18	246.6	39.5	2377	10	ADB98693 Human atr
19	246.6	39.5	2377	10	ADB84909 Farnesyl
20	246.6	39.5	2377	10	ADB82489 Human DNA

C 21	246.6	39.5	2559	2	AAZ09236	Az09236 Human E3
C 22	246.6	39.5	2678	12	ADB77069	Ade77069 Human CDN
C 23	246.6	39.5	3410	12	ADQ85919	Adq85919 Human tum
C 24	246.6	39.5	3410	12	ADQ84071	Adq84071 Human tum
C 25	246.6	39.5	3745	8	ACF34485	Acf34485 Gene enco
C 26	246.6	39.5	5372	2	AAZ09235	Az09235 Human E3
C 27	210	33.7	3288	13	ACN40036	Acn40036 Tumour-as
C 28	203.6	32.6	3475	4	AAD10511	Adi10511 Human ubi
C 29	203.6	32.6	3475	8	AAD50456	Adi50456 Nedd-4-1i
C 30	203.6	32.6	3475	12	ADQ19347	Adq19347 Human sof
C 31	203.6	32.6	3476	2	AAT95698	Aat95698 Novel hum
C 32	203.6	32.6	3476	10	ADB49241	Adb49241 Novel hum
C 33	168.8	27.1	4573	12	ADQ23594	Adq23594 Human sof
C 34	168.8	27.1	4573	12	ADQ23375	Adq23375 Human sof
C 35	135.2	21.7	4307	4	ABL06079	Abi06079 Drosophil
C 36	135.2	21.7	4390	4	ABL05629	Abi05628 Drosophil
C 37	111.8	17.9	2351	2	AAZ30406	Aax30406 DNA encod
C 38	111.8	17.9	2351	10	ADB47742	Adb47742 Novel hum
C 39	111.8	17.9	2351	12	ADJ55297	Adj55297 Novel hum
C 40	95.6	15.3	3663	12	ADQ64872	Adq64872 Novel hum
C 41	91.2	14.6	8354	4	ABL06078	Abi06078 Drosophil
C 42	91.2	14.6	8653	4	ABL05628	Abi05628 Drosophil
C 43	88.4	14.2	349	5	AAS80663	Aas80663 DNA encod
C 44	59.2	9.5	1863	5	AAS80665	Aas80665 DNA encod
C 45	56	9.0	2661	10	ADB53839	Adb53839 Primary r

ALIGNMENTS

RESULT 1

ADS00022

ID ADS00022 standard; DNA; 624 BP.

AC ADS00022;

DT 16-DEC-2004 (first entry)

DE Human WWP1 antisense fragment.

KW ds; gene; apoptosis-related disease; human; WWP1; apoptosis; cancer.

OS Homo sapiens.

PN US2004191220-A1.

PD 30-SEP-2004.

PF 11-JUL-2003; 2003US-00618408.

PR 11-JUL-2002; 2002US-0395358P.

PA (EINA/) EINAT P.

PA (DEIS/) DEISS L.

PA (MAYA/) MAYA R.

PI Einat P, Deiss L, Maya R;

PI WPI; 2004-698655/68.

PT Use of an inhibitor of the WWP1 polypeptide for treating an apoptosis-related disease in a subject such as cancer, and for preparing a medicament for treating the apoptosis-related disease.

PS Claim 4; SEQ ID NO 3; 32pp; English.

XX The invention relates to a method of treating an apoptosis-related disease in a subject which comprises administering to the subject a therapeutic amount of an inhibitor of the WWP1 polypeptide to inhibit WWP1 so as to thus treat the subject. The method is useful for treating an apoptosis-related disease, i.e. cancer. Inhibitor of human WWP1 polypeptide is useful in preparing a medicament for treating cancer. The present sequence represents the human WWP1 antisense fragment.

XX SQ Sequence 624 BP; 132 A; 150 C; 130 G; 204 T; 0 U; 8 Other;

Query Match 98.7%; Score 616; DB 13; Length 624;
Best Local Similarity 100.0%; Pred. No. 2.9e-184;
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCGCGGTTTAAAGCATTTGTGACCCCTTTAGTNCAGATGACTTCCATTGCGAGGAT 60
DB 1 CGCGCGGTTTAAAGCATTTGTGACCCCTTTAGTNCAGATGACTTCCATTGCGAGGAT 60

QY 61 CTTTGAATGTTGTTCTTGTGTTATGATCAACAAGTACCTTTACAGAGTAT 120
DB 61 CTTTGAATGTTGTTCTTGTGTTATGATCAACAAGTACCTTTACAGAGTAT 120

QY 121 ATCTAAATTTCCAGCTTCTGCGAGGGGTTCTTCAATCTGTGAAGCTTGTGAT 180
DB 121 ATCTAAATTTCCAGCTTCTGCGAGGGGTTCTTCAATCTGTGAAGCTTGTGAT 180

QY 181 CTTCCACTGGGTTGTTTTGTTATGATCAACAAGTAAACCTCTGTGTTGAATCCA 240
DB 181 CTTCCACTGGGTTGTTTTGTTATGATCAACAAGTAAACCTCTGTGTTGAATCCA 240

QY 241 CTTCTTTTCCAGCTGCTGGGAAAGTCCATTAAGGGTCAATTTCTGCGAGTAAACATTG 300
DB 241 CTTCTTTTCCAGCTGCTGGGAAAGTCCATTAAGGGTCAATTTCTGCGAGTAAACATTG 300

QY 301 AAGCGAATAGAGTATGTTGTTGTTAACTGTTGCATAGCTCCCTGCAATTTGGTCCGCT 360
DB 301 AAGCGAATAGAGTATGTTGTTGTTAACTGTTGCATAGCTCCCTGCAATTTGGTCCGCT 360

QY 361 GAGATGCCACTGNTCAAAATTCGAGACAGATTCATGTTAGGCGCTGCGAGTTGTTG 420
DB 361 GAGATGCCACTGNTCAAAATTCGAGACAGATTCATGTTAGGCGCTGCGAGTTGTTG 420

QY 421 TTTCTGGTTATGATACACATAATAAATCTCTACGATCATCAATCTTTTCCCAAC 480
DB 421 TTTCTGGTTATGATACACATAATAAATCTCTACGATCATCAATCTTTTCCCAAC 480

QY 481 CTGGAGTAAAGGTTGTGGTCTCTCCCATGTTGGTAGTTCAGCATTTATGANCACATAAT 540
DB 481 CTGGAGTAAAGGTTGTGGTCTCTCCCATGTTGGTAGTTCAGCATTTATGANCACATAAT 540

QY 541 ACGCTCTACCATGAGGATCTTTTCTTTGTTCCACCCCTGATGCGCAANGTCTGTGTTG 600
DB 541 ACGCTCTACCATGAGGATCTTTTCTTTGTTCCACCCCTGATGCGCAANGTCTGTGTTG 600

QY 601 GCATTTCCAGACTGCTGCCGNACA 624
DB 601 GCATTTCCAGACTGCTGCCGNACA 624

RESULT 2
AAT95697/c
ID AAT95697 standard; DNA; 2052 BP.
XX AC AAT95697;
XX DT 23-APR-1998 (first entry)
XX DE Novel human gene, designated WWP1.
XX KW Peptide recognition unit; YAP WW domain binding protein; WBP-1; WBP-2;
XX NW WW domain; cell signalling; growth regulation; cytoskeleton organisation;
XX NW targeted drug screening; modulator; WW domain interaction; ss.
XX OS Homo sapiens.
XX PN WO9737223-A1.
XX PD 09-OCT-1997.
XX PF 03-APR-1997; 97WO-US0005547.

XX PR 03-APR-1996; 96US-00630916.
XX PA (CYTO-) CYTOGEN CORP.
XX PA (UYNC-) UNIV NORTH CAROLINA.
XX PI Pirozzi G, Kay BK, Fowlkes DM;
XX DR WPI; 1997-503234/46.
XX DR P-PSDB; AAW36794.

XX PT Identifying cell signalling and growth regulatory polypeptides by
XX PT reaction with multivalent recognition complex - polypeptides are useful
XX PT in targetted drug selection.

XX PS Claim 68; Fig 16 A-B; 220pp; English.

XX CC The present DNA sequence encodes a novel protein WWP1. The WWP1 gene was
XX CC identified and isolated from human bone marrow and brain cDNA libraries,
XX CC using peptides AAW38103-05. These peptide recognition units are based on
XX CC the sequences of WW domain binding domains of the YAP WW domain binding
XX CC proteins WBP-1 and WBP-2. The WW domain is a small functional domain
XX CC found in a large number of proteins from a variety of species including
XX CC humans, nematodes and yeast. Its name is derived from the observation
XX CC that two tryptophan residues, one in the amino terminal portion of the WW
XX CC domain and one in the carboxyl terminal portion, are conserved. Most
XX CC proteins containing WW domains have a function involving cell signalling
XX CC and growth regulation or the organisation of the cytoskeleton.
XX CC Polypeptides containing a WW domain are identified by treating a
XX CC multivalent recognition unit complex that has selective binding affinity
XX CC for a WW domain, with many polypeptides and identifying those with
XX CC selective affinity for the complex. Proteins containing WW domains are
XX CC used for targeted drug screening, i.e. to identify potential modulators
XX CC of specific WW domain interactions. The valency of the recognition unit
XX CC is important in determining specificity of interaction with WW domains.
XX CC In multivalent form specificity is relaxed, but not lost, so proteins
XX CC containing WW domains similar, but not identical, to the sequence of the
XX CC peptides' target WW can be detected, including new polypeptides

SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 2; Length 2052;
Best Local Similarity 97.3%; Pred. No. 3e-168;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 86
DB 1044 CCTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 985

QY 87 TGATCAACAAGTACCTTACACCTTCCAGAGTATATCTAAATTTCCAGCTTCTGCGAGG 146
DB 984 TGATCAACAAGTACCTTACACCTTCCAGAGTATATCTAAATTTCCAGCTTCTGCGAGG 925

QY 147 GGTTCCTTATCTGTAAGCTTGAATCTTCCCATCTTCCCATCTGCGAGTATCTGTTGTTA 206
DB 924 GGTTCCTTATCTGTAAGCTTGAATCTTCCCATCTTCCCATCTGCGAGTATCTGTTGTTA 865

QY 207 TGATTCACAAGTAAACCTCTGCTGTAATCCACTCTTTTCCAGCTGCGGCAAA 266
DB 864 TGATTCACAAGTAAACCTCTGCTGTAATCCACTCTTTTCCAGCTGCGGCAAA 805

QY 267 GGTTCATAGGGTCAATTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326
DB 804 GGTTCATAGGGTCAATTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 745

QY 327 AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGTCAAAATTTCGG 386
DB 744 AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGTCAAAATTTCGG 685

QY 387 ACAGATTCATGTTAGGCGCTGCGACCGTTGTTGTTGTTGTTGTTATGATCCACATAATAA 446
DB 684 ACAGATTCATGTTAGGCGCTGCGACCGTTGTTGTTGTTGTTGTTATGATCCACATAATAA 625

QY 447 ACTCCTTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC 506
 DB 624 ACTCTTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC 565
 QY 507 CATGTGTAGTTCGAGCATATGACCACATATACGCTCTACCATGAGGATCTTTCTT 566
 DB 564 CATGTGTAGTTCGAGCATATGACCACATATAGTTCATCATGAGGATCTTTCTT 505
 QY 567 TGTTCACACCCCTGATGGCAANGTCTGTGTGGCAATCCAGACTGTGCCGNACA 624
 DB 504 TGTTCACA-CCCTGATGGCAAGTTCGTGTGGCAATCCAGACTGTGCCGTACA 448

RESULT 3
 AAD10504/c
 ID AAD10504 standard; DNA; 2052 BP.
 XX
 AC AAD10504;
 XX
 DT 24-SEP-2001 (first entry)
 XX
 DE Human ubiquitin protein ligase WWP1 DNA.
 XX
 KW Human; ubiquitin protein ligase; WWP1; antitumour; antiinflammatory;
 XX therapy; infection; inflammation; tumour; ds.
 OS Homo sapiens.
 PH Key Location/Qualifiers
 FT CDS 2..2050
 FT /*tag= a
 FT /product= "Human WWP1 protein"
 FT /note= "CDS does not include start and stop codon"
 FT /partial
 XX
 FN US6258601-B1.
 XX
 PD 10-JUL-2001.
 XX
 PF 07-SEP-2000; 2000US-00657481.
 XX
 PR 07-SEP-2000; 2000US-00657481.
 XX
 PA (ISIS-) ISIS PHARM INC.
 XX
 PI Monia BP, Cowesert LM;
 XX
 DR WPI; 2001-450370/48.
 DR P-PSDB; AAE05494.
 XX
 PT Antisense compounds capable of modulating expression of ubiquitin protein
 PT ligases WWP1 and WWP2, useful for diagnosis, prophylaxis and treatment of
 PT diseases e.g. infection, inflammation or tumors.
 XX
 PS Claim 1; Col 53-58; 47pp; English.
 XX
 CC The present invention relates to compounds, particularly antisense
 CC oligonucleotides, which are targeted to nucleic acids encoding ubiquitin
 CC protein ligases WWP1 and WWP2. The antisense oligonucleotides modulate
 CC the expression of WWP1 and WWP2. The antisense oligonucleotides are
 CC useful for inhibiting the expression of ubiquitin protein ligases WWP1
 CC and WWP2 in cells or tissues in vitro. The oligonucleotides are useful
 CC for diagnosing, treating diseases associated with the expression of
 CC ubiquitin protein ligases WWP1 and WWP2 and for prophylaxis e.g. to
 CC prevent or delay infection, inflammation or tumour formation and as a
 CC research reagent. The present sequence is a DNA encoding human ubiquitin
 CC protein ligase WWP1
 XX
 SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 4; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTTGCCAGGATCTTTGAATGTTGTTCTTGTGTTA 86
 DB 1044 CTTTATAGTTACAGATGACTTCCCATTTGCCAGGATCTTTGAATGTTGTTCTTGTGTTA 985
 QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGACCTTCTGGCAGG 146
 DB 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGACCTTCTGGCAGG 925
 QY 147 GGTTCCTTCATCTGTAAAGCCTTGAGTTCCTTGGATCTTCCACATGGGTGTTTGTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAAGCCTTGAGTTCCTTGGATCTTCCACATGGGTGTTTGTGTTA 865
 QY 207 TGATCAACAAAGTAAACCCCTGCTGTGATCCACTCTTTTCCAGACCTGTTGCAAA 266
 DB 864 TGATCAACAAAGTAAACCCCTGCTGTGATCCACTCTTTTCCAGACCTGTTGCAAA 805
 QY 267 GGTCCATAAGGGTCATTTTCTGCAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 326
 DB 804 GGTCCATAAGGGTCATTTTCTGCAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 745
 QY 327 AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTCGG 386
 DB 744 AACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTGCCACTGNTCAAAATTCGG 685
 QY 387 ACAGATTCATAGTAGGCGCTGCCACGTTGTTGTTCTGCTGTATGATCCACATATAA 446
 DB 684 ACAGATTCATAGTAGGCGCTGCCACGTTGTTGTTCTGCTGTATGATCCACATATAA 625
 QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC 506
 DB 624 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAAGTGTGTCTCTCC 565
 QY 507 CATGTGTAGTTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 DB 564 CATGTGTAGTTCGAGCATATGANCACATATAGCTCTACCATGAGGATCTTTTCTT 505
 QY 567 TGTTCACACCCCTGATGGCAANGTTCGTGTGGCAATCCAGACTGTGCCGNACA 624
 DB 504 TGTTCACA-CCCTGATGGCAAGTTCGTGTGGCAATCCAGACTGTGCCGTACA 448

RESULT 4
 ABZ34944/c
 ID ABZ34944 standard; cDNA; 2052 BP.
 XX
 AC ABZ34944;
 XX
 DT 05-FEB-2003 (first entry)
 XX
 DE Human gene expression profile polynucleotide SEQ ID NO 56.
 XX
 KW Human; artery; endothelium; umbilical; vein; aorta; pulmonary artery;
 KW bronchial epithelium; prostate; muscle; lung fibroblast; osteoblast;
 KW tumour; microarray; genome mapping; antibiotic; antiviral; antifungal;
 KW gene expression; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200274979-A2.
 XX
 PD 26-SEP-2002.
 XX
 PF 20-MAR-2002; 2002WO-US008456.
 XX
 PR 20-MAR-2001; 2001US-0276947P.
 XX
 PA (ORTH) ORTHO CLINICAL DIAGNOSTICS INC.
 XX
 PI Wan J, Wang Y;
 XX
 DR WPI; 2002-740862/80.
 XX

PT New gene expression profile generated from primary, endothelial, epithelial, and muscle cell types, useful for identifying disease pathologies involving alterations of gene expression, e.g. cancer.

PS

XX Claim 3; Page 256; 850pp; English.

CC The invention relates to a gene expression profile comprising one or more CC genes (ABZ34889-ABZ35692) and generated from a cell type. The cell type CC is a coronary artery endothelium, umbilical artery or vein endothelium, CC aortic endothelium, dermal microvascular endothelium, pulmonary artery CC endothelium, myometrium microvascular endothelium, keratinocyte CC epithelium, bronchial epithelium, mammary epithelium, prostate CC epithelium, renal cortical epithelium, renal proximal tubule epithelium, CC small airway epithelium, renal epithelium, umbilical artery smooth CC muscle, neonatal dermal fibroblast, pulmonary artery smooth muscle, CC dermal fibroblast, neural progenitor cells, skeletal muscle, astrocytes, CC aortic smooth muscle, mesangial cells, coronary artery smooth muscle, CC bronchial smooth muscle, uterine smooth muscle, lung fibroblast, CC osteoblasts or prostate stromal cell. The gene expression profile is used CC for determining the level of RNA expression for a sample, determining the CC phenotype of a cell and distinguishing cell types. The gene or a protein CC expression profile is useful in identifying disease pathologies involving CC alterations of gene expression. The assessment of expression profiles may CC provide meaningful information with respect to tumour type and stage, CC treatment methods, and prognosis. The gene or protein expression profile CC may also be used for creating microarrays. The microarray is useful for CC genetic and physical mapping of genomes, DNA sequencing, genetic or CC medical diagnosis, genotyping of organisms, confirming cell or tissue CC identifications and in identifying promising antibiotics, antiviral or CC antifungal agents

XX

SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 6; Length 2052;
Best Local Similarity 97.3%; Pred. No. 3e-168; Indels 1; Gaps 1;
Matches 582; Conservative 0; Mismatches 15;

QY 27 CCCTTAGTGCAGATGCTCCCATGCGAGGATCTTTGAATGTTGTTCTTGTGTTA 86
DB 1044 CCTTAGTTACAGATGACTCCCATGCGAGGATCTTTGAATGTTGTTCTTGTGTTA 985

QY 87 TGATCAACAAGTACCTTACACCTTACGAGTATATCTAATTTCCAGCGCTTCGCGAGG 146
DB 984 TGATCAACAAGTACCTTACACCTTACGAGTATATCTAATTTCCAGCGCTTCGCGAGG 925

QY 147 GGTCTTTCATCTGTAAGCCTTCAGTCTTTGGAATCTTCCACATGGGTGTTTGTGTTA 206
DB 924 GGTCTTTCATCTGTAAGCCTTCAGTCTTTGGAATCTTCCACATGGGTGTTTGTGTTA 865

QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTCAGCGCTGTTGCAAA 266
DB 864 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTTCAGCGCTGTTGCAAA 805

QY 267 GGTCCATAGGGTCAATTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
DB 804 GGTCCATAGGGTCAATTTCTGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

QY 327 AACTGTTGCTAGCTCCCTGCAATTCGTCCTCGGTGAGATGTCACATGNTCAAAATTCGG 386
DB 744 AACTGTTGCTAGCTCCCTGCAATTCGTCCTCGGTGAGATGTCACATGNTTCGAGTTTCGG 685

QY 387 ACAGATTCCATGTTAGCGCGCTGCCAGTTGTTGTTGTTGTTATGATCCACATATAA 446
DB 684 ACAGATTCCATGTTAGCGCGCTGCCAGTTGTTGTTGTTGTTATGATCCACATATAA 625

QY 447 ACTCCTCTAGCATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTCCTCC 506
DB 624 ACTCCTCTAGCATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTCCTCC 565

QY 507 CATGTGGTGTGTCGAGCATTTATGANCACATATACGCTCTACCATGAGGATCTTTTCTT 566
DB 564 CATGTGGTGTGTCGAGCATTTATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT 505

QY 567 TGTGCCACCCCTGATGGCAAGNTTCTGTGTGGCAATTCGCCAGACTGCTGCCGNACA 624
DB 504 TGTGCCA-CCCTGATGGCAAGGTTTCTGTGTGGCAATTCGCCAGACTGCTGCCGTACA 448

RESULT 5
AAD50455/c
ID AAD50455 standard; DNA; 2052 BP.
XX
AC AAD50455;
XX
DT 24-MAR-2003 (first entry)
XX
DE Nedd-4-like ubiquitin-protein ligase WWP1 DNA.
XX
DE Viral infection; lymphosarcoma; human immunodeficiency virus; hepatitis;
XX poliomylitis; HIV; measles; protein therapy; enzyme; gene;
KW Nedd-4-like ubiquitin-protein ligase; ds.
XX
OS Unidentified.
XX
FH Key Location/Qualifiers
FT CDS 2..2052
FT /tag= a
FT /product= "Nedd-4-like ubiquitin-protein ligase WWP1"
FT /transl_except= (pos:2051..2052, aa:Ser)
FT /notes= "CDS does not include start and stop codon"
FT /partial
XX
PN WO200290549-A2.
XX
PD 14-NOV-2002.
XX
PF 12-MAR-2002; 2002WO-1B002106.
XX
PR 12-MAR-2001; 2001US-0275224P.
PR 31-JUL-2001; 2001US-0308958P.
PR 07-DEC-2001; 2001US-0340170P.
XX
PA (PROT-) PROTEOLOGICS LTD.
XX
PI Greener T, Moskowitz H, Reiss Y, Alroy I;
XX
DR WPI; 2003-111976/10.
DR P-PSDB; AAE32721.
XX
XX New protein complex comprising HECT-RC1, viral maturation scaffolding
PT protein (VMSF), and/or HIV gag protein, useful for treating viral
PT infections, such lymphosarcoma, HIV, hepatitis, poliomyelitis, measles,
PT or Ebola.
XX
PS Disclosure; Fig 3; 150pp; English.
XX
CC The invention relates to a method for modulation of viral maturation. The
CC invention also provides an isolated protein complex comprising a HECT-
CC RCC1 selected from HECT-WW, HECT-RC1, Gag protein, Gag late domain, P13,
CC actin, myosin, Hsp70, Hsp90, STAM1, STAM2, VHS-UM,
CC GTPase, E2 enzyme, tsq101, cullin, HEC1, HEC2, HEC3, Nedd4-like
CC protein or clathrin. The complexes, proteins, antibodies and methods are
CC useful for treating viral infections, such as lymphosarcoma, human
CC immunodeficiency virus (HIV), hepatitis, poliomyelitis, measles, or Ebola
CC and for inhibiting budding in a subject. They are also useful in
CC diagnostic assays for determining whether a cell is infected with a virus
CC and for characterising the nature, progression and/or infectivity of the
CC infection. The invention is also useful in protein therapy. The present
CC sequence is Nedd-4-like ubiquitin-protein ligase DNA used to illustrate
CC the method of the invention
XX
SQ Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;

Query Match 90.7%; Score 566.2; DB 8; Length 2052;
Best Local Similarity 97.3%; Pred. No. 3e-168; Indels 1; Gaps 1;
Matches 582; Conservative 0; Mismatches 15;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 86
 DB 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 985
 QY 87 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 146
 DB 984 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 925
 QY 147 GGTTCCTTCATCTGTAAGCCTTGAAGTCTTCCCACTGGGTTGTTTGTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAGCCTTGAAGTCTTCCCACTGGGTTGTTTGTGTTA 865
 QY 207 TGATTCACAAAGTAAACCTCTGTTGTTGAATCCACTCTTTTTCACAGCCCTTCGGCAAA 266
 DB 864 TGATTCACAAAGTAAACCTCTGTTGTTGAATCCACTCTTTTTCACAGCCCTTCGGCAAA 805
 QY 267 GGTTCATTAAGGTCATTTCTGACGCTAACATTTGAAGCCGAATAGAGTATCGTTGTTA 326
 DB 804 GGTTCATTAAGGTCATTTCTGACGCTAACATTTGAAGCCGAATAGAGTATCGTTGTTA 745
 QY 327 AACTGTTGCTAGTCTCCCTGCAATTTGCTCGCTGAGATGCCACTGNTCAAAATTCGG 386
 DB 744 AACTGTTGCTAGTCTCCCTGCAATTTGCTCGCTGAGATGCCACTGNTCAAAATTCGG 685
 QY 387 ACAGATTCATGCTAGGCGCTGCGCTGAGATGCCACTGNTCAAAATTCGG 446
 DB 684 ACAGATTCATGCTAGGCGCTGCGCTGAGATGCCACTGNTCAAAATTCGG 625
 QY 447 ACTCTCTACATCATCACTCTTCTTCCCACTGGAGTAAAGTTGGTCTCTCC 506
 DB 624 ACTCTCTACATCATCACTCTTCTTCCCACTGGAGTAAAGTTGGTCTCTCC 565
 QY 507 CATGTGCTAGTTCGAGCATTTATGATCCACATATAGAGTTCTACCATGAGGATCTTTCTT 566
 DB 564 CATGTGCTAGTTCGAGCATTTATGATCCACATATAGAGTTCTACCATGAGGATCTTTCTT 505
 QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGTACA 624
 DB 504 TGTTCCTCA-CCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGTACA 448

RESULT 6
 ADB49239/c
 ID ADB49239 standard; DNA; 2052 BP.
 XX AC ADB49239;
 XX AC ADB49239;
 XX DT 04-DEC-2003 (first entry)
 XX DE Novel human WWP (WW binding protein) 1 DNA.
 XX KW WW domain; drug candidate screening; drug discovery; drug modification;
 XX KW drug refinement; immunogen; WWP1; WW binding protein 1; gene; ds; human.
 XX OS Homo sapiens.
 XX XX US2003077577-A1.
 XX XX 24-APR-2003.
 XX XX 28-JUN-2002; 2002US-00185050.
 XX XX 03-APR-1996; 96US-00630916.
 XX XX 03-APR-1997; 97US-00826516.
 XX XX (PIROZZI) PIROZZI G.
 XX XX (KAYB/) KAY B K.
 XX XX (FOWL/) FOWLKES D M.
 XX XX Pirozzi G, Kay BK, Fowlkes DM;
 XX XX WPI; 2003-635075/60.

DR P-PSDB; ADB49240.
 XX Novel purified polypeptide comprising WW domain, useful for drug
 PT discovery, modification and refinement, for discovering polypeptides
 PT involved in pharmacological activities, or as an immunogen to generate
 PT antibodies.
 XX Claim 68; Fig 16; 133pp; English.
 CC The invention describes a purified polypeptide (I) comprising a WW domain
 CC which has a sequence (S1) selected from 11 sequences fully defined in the
 CC specification, a sequence (S2) selected from 48 sequences fully defined
 CC in the specification or a sequence (S3) comprising 683, 906, 224 or 725
 CC amino acids fully defined in the specification. (I) is useful for
 CC screening a potential drug candidate, by allowing (I) to come into
 CC contact with at least one recognition unit having a selective affinity
 CC for the WW domain in (I), in the presence of an amount of a potential
 CC drug candidate, such that (I) and the recognition unit are capable of
 CC interacting when brought into contact with one another in the absence of
 CC the drug candidate, and determining the effect, if any, of the presence
 CC of the amount of the drug candidate on the interaction of (I) with the
 CC recognition unit. (I) is useful for drug discovery, modification and
 CC refinement, for discovering polypeptides involved in pharmacological
 CC activities, or as an immunogen to generate antibodies. This sequence
 CC encodes a novel human WW binding protein WWP1.
 XX Sequence 2052 BP; 636 A; 405 C; 447 G; 564 T; 0 U; 0 Other;
 SQ
 Query Match 90.7%; Score 566.2; DB 10; Length 2052;
 Best Local Similarity 97.3%; Pred. No. 3e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 86
 DB 1044 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTCTTGTGTTA 985
 QY 87 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 146
 DB 984 TGATCAACAAGTACCTTACACCTTCAGAGTATATCTAATTTCCAGCCCTTCGGCAGG 925
 QY 147 GGTTCCTTCATCTGTAAGCCTTGAAGTCTTCCCACTGGGTTGTTTGTGTTA 206
 DB 924 GGTTCCTTCATCTGTAAGCCTTGAAGTCTTCCCACTGGGTTGTTTGTGTTA 865
 QY 207 TGATTCACAAAGTAAACCTCTGTTGTTGAATCCACTCTTTTTCACAGCCCTTCGGCAAA 266
 DB 864 TGATTCACAAAGTAAACCTCTGTTGTTGAATCCACTCTTTTTCACAGCCCTTCGGCAAA 805
 QY 267 GGTTCATTAAGGTCATTTCTGACGCTAACATTTGAAGCCGAATAGAGTATCGTTGTTA 326
 DB 804 GGTTCATTAAGGTCATTTCTGACGCTAACATTTGAAGCCGAATAGAGTATCGTTGTTA 745
 QY 327 AACTGTTGCTAGTCTCCCTGCAATTTGCTCGCTGAGATGCCACTGNTCAAAATTCGG 386
 DB 744 AACTGTTGCTAGTCTCCCTGCAATTTGCTCGCTGAGATGCCACTGNTCAAAATTCGG 685
 QY 387 ACAGATTCATGCTAGGCGCTGCGCTGAGATGCCACTGNTCAAAATTCGG 446
 DB 684 ACAGATTCATGCTAGGCGCTGCGCTGAGATGCCACTGNTCAAAATTCGG 625
 QY 447 ACTCTCTACATCATCACTCTTCTTCCCACTGGAGTAAAGTTGGTCTCTCC 506
 DB 624 ACTCTCTACATCATCACTCTTCTTCCCACTGGAGTAAAGTTGGTCTCTCC 565
 QY 507 CATGTGCTAGTTCGAGCATTTATGATCCACATATAGAGTTCTACCATGAGGATCTTTCTT 566
 DB 564 CATGTGCTAGTTCGAGCATTTATGATCCACATATAGAGTTCTACCATGAGGATCTTTCTT 505
 QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGTACA 624
 DB 504 TGTTCCTCA-CCCTGATGGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGTACA 448

RESULT 7
 ADC26288/c
 ID ADC26288 standard; DNA; 2731 BP.
 XX
 AC ADC26288;
 XX
 DT 18-DEC-2003 (first entry)
 XX
 D8 Human NOV27 DNA.
 XX
 KW NOV; cytostatic; metabolic disorder; immune; neurodegenerative;
 KW circulatory; haemopoietic; wasting; cancer; gene therapy; vaccine;
 KW transgenic; human; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN WO2003004687-A2.
 XX
 PD 16-JAN-2003.
 XX
 PF 03-JUL-2002; 2002WO-US021361.
 XX
 PR 05-JUL-2001; 2001US-0303046P.
 PR 09-JUL-2001; 2001US-0303828P.
 PR 09-JUL-2001; 2001US-0304016P.
 PR 11-JUL-2001; 2001US-0304502P.
 PR 13-JUL-2001; 2001US-0305262P.
 PR 16-JUL-2001; 2001US-0305673P.
 PR 17-JUL-2001; 2001US-0306085P.
 PR 24-JUL-2001; 2001US-0307536P.
 PR 27-JUL-2001; 2001US-0308228P.
 PR 30-JUL-2001; 2001US-0308877P.
 PR 01-AUG-2001; 2001US-0309255P.
 PR 17-AUG-2001; 2001US-0313328P.
 PR 12-SEP-2001; 2001US-0318711P.
 PR 19-SEP-2001; 2001US-0323380P.
 PR 21-SEP-2001; 2001US-0323969P.
 PR 04-JAN-2002; 2002US-0345023P.
 PR 04-JAN-2002; 2002US-0345038P.
 PR 28-FEB-2002; 2002US-0361172P.
 PR 01-MAR-2002; 2002US-0360814P.
 PR 01-MAR-2002; 2002US-0360830P.
 PR 01-MAR-2002; 2002US-0361133P.
 PR 01-MAR-2002; 2002US-0361147P.
 PR 05-MAR-2002; 2002US-0361677P.
 PR 02-APR-2002; 2002US-0363637P.
 PR 12-APR-2002; 2002US-0372326P.
 PR 16-APR-2002; 2002US-0372990P.
 PR 19-APR-2002; 2002US-0373881P.
 PR 19-APR-2002; 2002US-0373921P.
 PR 02-JUL-2002; 2002US-00188186.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 PI Anderson DW, Berghs C, Boldog FL, Burgess CE, Casman SJ;
 PI Catterton E, Edinger S, Eisen AJ, Ellerman K, Gerlach V, Gorman L;
 PI Guo X, Jeffers M, Kekuda R, Li L, Malyankar UM, Miller CE;
 PI Padigaru M, Patturajan M, Pena CE, Rastelli L, Shenoy S;
 PI Shimkets RA, Spaderna SK, Spytek KA, Stone DJ, Taupier RJ;
 PI Vernet CM, Voss EZ, Zhong M;
 XX
 WPI; 2003-221607/21.
 DR P-PSDB; ADC26289.
 DR
 XX
 PT New isolated NOVX polypeptide, useful for determining the presence of, or
 PT predisposition to a disease associated with altered levels of expression
 PT of the polypeptide, and for treating or preventing cancer.
 XX
 PS Claim 20; SEQ ID NO 113; 478pp; English.
 XX
 CC The invention relates to a novel isolated NOV polypeptide. The
 CC polypeptide of the invention demonstrates cytosstatic activity and may be
 CC used for determining the presence of, or predisposition to a disease

CC associated with altered levels of expression of the polypeptide,
 CC including metabolic disorders, immune disorders, neurodegenerative
 CC disorders, circulatory diseases, haemopoietic disorders, wasting diseases
 CC and cancer. The polypeptide may also be utilised during gene therapy
 CC procedures, vaccine development and transgenic animal production. The
 CC current sequence is that of the human NOV DNA of the invention.
 XX
 SQ Sequence 2731 BP; 896 A; 505 C; 599 G; 731 T; 0 U; 0 Other;
 Query Match 90.7%; Score 566.2; DB 10; Length 2731;
 Best Local Similarity 97.3%; Pred. No. 3.5e-168;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCTTTAGTNCAGATGACATCCCACTCCCAATGCGAGGATCTTTGAATGTTGTTCTTTGTTA 86
 DB 1620 CCTTTAGTTACAGATGACTTCCCAATGCGAGGATCTTTGAATGTTGTTCTTTGTTA 1561
 QY 87 TGATCAACAAAGTACCTTACACCTTCCAGATATATCTAATTTCCAGACCTTCTGCGAGG 146
 DB 1560 TGATCAACAAAGTACCTTACACCTTCCAGATATATCTAATTTCCAGACCTTCTGCGAGG 1501
 QY 147 GGTTCCTTCATCTGTAAAGCCTTGAGTTCTTGAATCTTCCACTGGGTTGTTTGTGTTA 206
 DB 1500 GGTTCCTTCATCTGTAAAGCCTTGAGTTCTTGAATCTTCCACTGGGTTGTTTGTGTTA 1441
 QY 207 TGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGACCTTCTGCGAGG 266
 DB 1440 TGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGACCTTCTGCGAGG 1381
 QY 267 GGTCCATAAGGGTCAATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326
 DB 1380 GGTCCATAAGGGTCAATTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 1321
 QY 327 AACTGTTGATAGTCCCTGCAATTTGGTTCCGCTGAGATTGCCACTGNTCAAAATNCGG 386
 DB 1320 AACTGTTGATAGTCCCTGCAATTTGGTTCCGCTGAGATTGCCACTGNTCAAAATNCGG 1261
 QY 387 ACAGATTCCATGCTAGGCGCTGCCAGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 446
 DB 1260 ACAGATTCCATGCTAGGCGCTGCCAGCTTGTGTTGTTGTTGTTGTTGTTGTTGTTA 1201
 QY 447 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 506
 DB 1200 ACTCCTCTACGATCATCACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 1141
 QY 507 CATGTGGTAGTTCGAGCATTATGACCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 DB 1140 CATGTGGTAGTTCGAGCATTATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT 1081
 QY 567 TGTTCCTCCACCTCGATGCAANGNTTCTGTGTTGGCATTTCCAGACTGCTGCCGNACA 624
 DB 1080 TGTTCCTCA-CCCTGATGCCAGGTTTCTGTGTTGCCATTTCCAGACTGCTGCCGTACA 1024
 RESULT 8
 ADS00020/c
 ID ADS00020 standard; DNA; 2793 BP.
 XX
 AC ADS00020;
 XX
 DT 16-DEC-2004 (first entry)
 XX
 DE Human WWP1 gene.
 DE
 KW ds; gene; apoptosis-related disease; human; WWP1; apoptosis; cancer.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..2793 /tag= a
 FT /product= "WWP1"
 FT
 XX

PN US2004191220-A1.
XX 30-SEP-2004.
XX 11-JUL-2003; 2003US-00618408.
XX 11-JUL-2002; 2002US-0395358P.
XX (EINA/) EINAT P.
XX (DEIS/) DEISS L.
XX (MAYA/) MAYA R.
XX Binat P, Deiss L, Maya R;
XX WPI; 2004-698655/68.
XX P-PSDB; ADS00021.
XX Use of an inhibitor of the WWP1 polypeptide for treating an apoptosis-
XX related disease in a subject such as cancer, and for preparing a
XX medicament for treating the apoptosis-related disease.
XX Disclosure; SEQ ID NO 1; 32pp; English.
XX The invention relates to a method of treating an apoptosis-related
XX disease in a subject which comprises administering to the subject a
XX therapeutic amount of an inhibitor of the WWP1 polypeptide to inhibit
XX WWP1 so as to thus treat the subject. The method is useful for treating
XX an apoptosis-related disease, i.e. cancer. Inhibitor of human WWP1
XX polypeptide is useful in preparing a medicament for treating cancer. The
XX present sequence represents the human WWP1 gene.
XX Sequence 2793 BP; 921 A; 523 C; 608 G; 741 T; 0 U; 0 Other;
Query Match 90.7%; Score 566.2; DB 13; Length 2793;
Best Local Similarity 97.3%; Pred. No. 3.5e-168;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCCTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTCTTGTTA 86
DB 1610 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTCTTGTTA 1551
QY 87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 146
DB 1550 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 1491
QY 147 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACTGGGTTGTTTGTGTTA 206
DB 1490 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACTGGGTTGTTTGTGTTA 1431
QY 207 TGATTCACAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCAGCCTGGTGCAAA 266
DB 1430 TGATTCACAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTTCAGCCTGGTGCAAA 1371
QY 267 GGTCCATAGGGTCATTTTCTGAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 326
DB 1370 GGTCCATAGGGTCATTTTCTGAGCTAAACATTGAAGCCGAATAGAGGTATCGTTGTTA 1311
QY 327 AACTGTTGCATAGCTCCCTCAATTTGTTCCGCTGAGATTGCCACTGNTCAAAATTCGG 386
DB 1310 AACTGTTGCATAGCTCCCTCAATTTGTTCCGCTGAGATTGCCACTGNTCAAAATTCGG 1251
QY 387 ACAGATTCCATGTTAGGCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATA 446
DB 1250 ACAGATTCCATGTTAGGCGCTGCGACGTTGTTGTTCTGTTGTTATGATCCACATAATA 1191
QY 447 ACTCCTCTAGCATCAACATCTTTCTTTTCCCAACCTGGAGGTAAAGTTGTTGTTCTCTCC 506
DB 1190 ACTCCTCTAGCATCAACATCTTTCTTTTCCCAACCTGGAGGTAAAGTTGTTGTTCTCTCC 1131
QY 507 CATGTGTAGTTCGAGCATTTATGNCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
DB 1130 CATGTGTAGTTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 1071

QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGNACA 624
DB 1070 TGTTCACA-CCCTGATGGCAAGGTTCTGTGTTGGCATTCCTCCAGACTGCTGCCGTACA 1014
RESULT 9
AAC90422/c
ID AAC90422 standard; cDNA; 2827 BP.
XX AAC90422;
XX 19-MAR-2001 (first entry)
XX Human clone 811a coding sequence.
XX Human; angiogenesis; cancer; Drosophila suppressor of deltex; Su(dx);
XX CADASIL; wound healing; rheumatoid arthritis; vascular disease;
XX arteriosclerosis; ss.
XX Homo sapiens.
XX WO200073329-A2.
XX 07-DEC-2000.
XX 23-MAY-2000; 2000WO-GB001990.
XX 26-MAY-1999; 99GB-00012132.
XX (UYMA-) UNIV VICTORIA MANCHESTER.
XX Baron M;
XX WPI; 2001-061509/07.
XX P-PSDB; AAB50048.
XX Use of homologs of Drosophila Notch regulator gene and encoded protein
XX products and antibodies in diagnosis and therapy of breast cancer,
XX angiogenesis and diseases associated with abnormal notch signaling.
XX Example; Page 28-30; 44pp; English.
XX The present invention relates to a human homolog of Drosophila suppressor
XX of deltex (Su(dx)) coding sequence and protein (see AAC90423 and
XX AAB50049). The human homologs are useful for in vitro diagnosis or
XX therapy of diseases such as angiogenesis, colon cancer, cervical cancer,
XX breast cancer, squamous adenocarcinoma, seminoma, melanoma, lung cancer,
XX dementia, cerebral autosomal dominant arteriopathy with sub-cortical
XX infarcts and leukoencephalopathy (CADASIL), wound healing, rheumatoid
XX arthritis, vascular diseases such as arteriosclerosis. The present
XX sequence is human clone 811a. This sequence has high homology to
XX Drosophila Su(dx). This sequence was used to screen a human breast tumour
XX cDNA library, resulting in identification of the human homolog of
XX AAC90423
XX Sequence 2827 BP; 917 A; 522 C; 578 G; 810 T; 0 U; 0 Other;
Query Match 90.7%; Score 566.2; DB 4; Length 2827;
Best Local Similarity 97.3%; Pred. No. 3.5e-168;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCCTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 86
DB 1043 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 984
QY 87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 146
DB 983 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCTTCGCGAGG 924
QY 147 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACTGGGTTGTTTGTGTTA 206
DB 923 GGTTCCTTCATCTGTAGGCTTGGATCTTGGATCTTCCACTGGGTTGTTTGTGTTA 864

QY 207 TGATTACAAAGTAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCTGGTGGCAA 266
Db |||||
863 TGATTACAAAGTAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCTGGTGGCAA 804
QY 267 GGTCCATAAGGGTCATTTCTGCAGCTAACATTGAAGCCGAATAGAGGATCGTTGGTTA 326
Db |||||
803 GGTCCATAAGGGTCATTTCTGCAGCTAACATTGAAGCCGAATAGAGGATCGTTGGTTA 744
QY 327 AACTGTTGCATAGCTCCCTGCAATTGGTTCGGTGGATTCGCACCTGNTCAAAATTCG 386
Db |||||
743 AACTGTTGCATAGCTCCCTGCAATTGGTTCGGTGGATTCGCACCTGNTCAAAATTCG 684
QY 387 ACAGATTCCATGTAGCGCGCTGCCAGTGTGTTGTTCTGGTGTATGATCCACATAATA 446
Db |||||
683 ACAGATTCCATGTAGCGCGCTGCCAGTGTGTTGTTCTGGTGTATGATCCACATAATA 624
QY 447 ACTCTCTACGATCATCACTCTTTTCCCACTGGAGGTAAGAGTTGGTCTCTCC 506
Db |||||
623 ACTCTCTACGATCATCACTCTTTTCCCACTGGAGGTAAGAGTTGGTCTCTCC 564
QY 507 CATGTGTTAGTTCGAGCATTTATGANCACACATAATAGCTCTACCATGAGGATCTTTTCT 566
Db |||||
563 CATGTGTTAGTTCGAGTATTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCT 504
QY 567 TGTTCACCCCTGTAGTGGCAANGTTCTGTGTTGGCATTCACAGACTGCTGCCGNACA 624
Db |||||
503 TGTTCACCA-CCCTGATGGCAAGTTTCTGTGTTGGCATTCACAGACTGCTGCCGTACA 447

RESULT 10
ABA93745/c
ID ABA93745 standard; cDNA; 3382 BP.

AC ABA93745;

XX 30-APR-2002 (first entry)

DE Human signal transduction cDNA clone tes3_11d21.

DE Human; foetal brain; foetal kidney; melanoma; testis; amygdala;
KW Gene therapy; ss.

XX Homo sapiens.

XX WO200198454-A2.

PN 27-DEC-2001.

PD 25-APR-2001; 2001WO-IB002050.

XX 25-APR-2000; 2000US-0199380P.

XX (GEHU-) GERMAN HUMAN GENOME PROJECT.

XX Wiemann S;

XX WPI; 2002-055860/07.

DR P-PSDB; ABB05708.

XX Human cDNA sequences and clones derived from human fetal brain, fetal
PT kidney, melanoma, testis and amygdala cDNA libraries, useful in genetic
PT screening and therapy.

PS Claim 1; Page 290-291; 611pp; English.

XX The present invention describes assemblages and computer readable media
CC comprising novel human cDNA sequences and clones derived from human
CC foetal brain, foetal kidney, melanoma, testis and amygdala cDNA
CC libraries. ABA93702 to ABA93766 represent human cDNA sequences from the
CC present invention which encode the proteins given in ABB05662 to
CC ABB05729. The human cDNA sequences and clones can be used in gene
CC therapy. The clones may be used in a variety of applications, for example
CC they may be used in profiling assays, for providing large arrays of human

CC genetic material for implementing large-scale screening strategies and
CC for treating diseases via gene therapy procedures
XX
SQ Sequence 3382 BP; 1115 A; 617 C; 705 G; 945 T; 0 U; 0 Other;
Query Match 90.7%; Score 566.2; DB 6; Length 3382;
Best Local Similarity 97.3%; Pred. No. 3.8e-168;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCCTTTAGTTCAGATGACCTTCCCATTTCCGAGATCTTTCAATGTTGTTCTTTGTTA 86
Db |||||
1620 CTTTAGTTACAGATGACTTCCCATTTCCGAGATCTTTCAATGTTGTTCTTTGTTA 1561
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 146
Db |||||
1560 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 1501
QY 147 GGTTCCTTCACTCTGTAAGCCTTGAGTCTTGGAATCTTCCACTGGGTTGTTTGTGTTA 206
Db |||||
1500 GGTTCCTTCACTCTGTAAGCCTTGAGTCTTGGAATCTTCCACTGGGTTGTTTGTGTTA 1441
QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCTGGTGGCAA 266
Db |||||
1440 TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCAGCCTGGTGGCAA 1381
QY 267 GGTTCATTAAGGGTCATTTTCTGCAGCTAACATTGAAGCCGAATAGAGGATCGTTGGTTA 326
Db |||||
1380 GGTTCATTAAGGGTCATTTTCTGCAGCTAACATTGAAGCCGAATAGAGGATCGTTGGTTA 1321
QY 327 AACTGTTGCATAGCTCCCTGCAATTGGTTCGGTGGATTCGCACCTGNTCAAAATTCG 386
Db |||||
1320 AACTGTTGCATAGCTCCCTGCAATTGGTTCGGTGGATTCGCACCTGNTCAAAATTCG 1261
QY 387 ACAGATTCCATGTAGCGCGCTGCCAGTGTGTTGTTCTGCTGTATGATCCACATAATA 446
Db |||||
1260 ACAGATTCCATGTAGCGCGCTGCCAGTGTGTTGTTCTGCTGTATGATCCACATAATA 1201
QY 447 ACTCTCTACGATCATCACTCTTTTCCCACTGGAGGTAAGAGTTGGTCTCTCC 506
Db |||||
1200 ACTCTCTACGATCATCACTCTTTTCCCACTGGAGGTAAGAGTTGGTCTCTCC 1141
QY 507 CATGTGTTAGTTCGAGCATTTATGANCACACATAATAGCTCTACCATGAGGATCTTTTCT 566
Db |||||
1140 CATGTGTTAGTTCGAGTATTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCT 1081
QY 567 TGTTCACCCCTGTAGTGGCAANGTTCTGTGTTGGCATTCACAGACTGCTGCCGNACA 624
Db |||||
1080 TGTTCACCA-CCCTGATGGCAAGTTTCTGTGTTGGCATTCACAGACTGCTGCCGTACA 1024

RESULT 11

ACN3827/c

ID ACN3827 standard; cDNA; 943 BP.

XX ACN3827;

XX 18-NOV-2004 (first entry)

DE Tumour-associated antigenic target (TAT) cDNA DNA325120, SEQ ID NO:2583.

XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
KW gene therapy; cytostatic; gene; ss.

XX Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

147 GGTCTCTCATCTGTAGCCCTTGAGTCTTGATCTTCCACCTGGGTTGTTTGTGTTA 206
1592 GGTCTCTCATCTGTAGCCCTTGAGTCTTGATCTTCCACCTGGGTTGTTTGTGTTA 1533
207 TGATTCAAAAGTAAACCCCTGTCTGTTGAATCACTCTTTTCCAGACCTGTGTGCAAA 266
1532 TGATTCAAAAGTAAACCCCTGTCTGTTGAATCACTCTTTTCCAGACCTGTGTGCAAA 1473
267 GGTCCATAAGGGTCATTTCTGCAGCTAAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326
1472 GGTCCATAAGGGTCATTTCTGCAGCTAAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 1413
327 AACTGTTGTCATAGCTCCCTGCAATTTGTTCCGCTGAGATTTGCCACTGNTCAAAATTCOGG 386
1412 AACTGTTGTCATAGCTCCCTGCAATTTGTTCCGCTGAGATTTGCCACTGTTCAAAATTCOGG 1353
387 ACAGATTCATGTTAGGCCCTGCGCAGTTGTTGTTCTGTTGATATGATCCACATAATA 446
1352 ACAGATTCATGTTAGGCCCTGCGCAGTTGTTGTTCTGTTGATATGATCCACATAATA 1293
447 ACTCCTCTACGATCATCACTCTTCTTCCACCTGGAGTAAAGGTTGTTGTTCTCTCC 506
1292 ACTCTCTACGATCATCACTCTTCTTCCACCTGGAGTAAAGGTTGTTGTTCTCTCC 1233
507 CATGTGTGATGTTGAGCAATATGANCACACATAATAGCTCTACCATGAGGATCTTTTCTT 566
1232 CATGTGTGATGTTGAGCAATATGATCCACATAATAGTTCTACCATGAGGATCTTTTCTT 1173
567 TGTTCACACCCCTGATGGCAANGTTCTGTGTTGGCATTTCCAGACTGCTGCGCGTACA 624
1172 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTTGGCATTTCCAGACTGCTGCGCGTACA 1116

RESULT 13
ADP84403/c
ID ADP84403 standard; DNA; 2305 BP.
XX AC ADP84403;
XX
DT 09-SEP-2004 (first entry)
XX
DE Human breast-specific protein coding sequence #2.
XX human; breast-specific protein; breast cancer; gene; ds.
XX Homo sapiens.
XX WO2004053077-A2.
XX
PD 24-JUN-2004.
XX
PF 05-DEC-2003; 2003WO-US038815.
XX
PR 05-DEC-2002; 2002US-0431123P.
XX
PA (DIAD-) DIADEXUS INC.
XX
PI Macina RA, Turner LR, Sun Y, Chen H, Rodriguez M;
XX WPI; 2004-468848/44.
DR P-PSDB; ADP84498, ADP84499.
DR
PT New breast specific nucleic acid molecules and polypeptides useful for
XX diagnosing, preventing or treating breast cancer, for producing
PT transgenic animals or cells, or for research purposes.
XX
PS Claim 1; SEQ ID NO 2; 521pp; English.
XX
CC The invention comprises the amino acid and coding sequences of human
CC breast-specific proteins. The DNA and protein sequences of the invention
CC are useful for the diagnosis, treatment and prevention of breast cancer.
CC The present DNA sequence encodes a human breast-specific protein of the
CC invention.

XX SQ Sequence 2305 BP; 712 A; 571 C; 545 G; 476 T; 0 U; 1 Other;
Query Match 73.2%; Score 456.8; DB 12; Length 2305;
Best Local Similarity 96.0%; Pred. No. 1.4e-133;
Matches 486; Conservative 0; Mismatches 18; Indels 2; Gaps 2;
QY 120 TATCTAATTTCCAGCCTTCTGSC-AGGGGTTCTTTCATCTGTAAAGCCTTGAGTCTTGG 178
Db 1868 TCTAAATTTGCGAGCCTTCTGCCAAGGGTTTCTTCAATCTGTAAAGCCTTGAGTCTTGG 1809
QY 179 ATCTTCCCACTGGGTTGTTTGTGTTATGATTCACAAAAGTAAACCCCTGTCTGTGAATC 238
Db 1808 ATCTTCCCACTGGGTTGTTTGTGTTATGATTCACAAAAGTAAACCCCTGTCTGTGAATC 1749
QY 239 CACTCTTTTTCAGCCTGTTGGCAAGGTTCATAGGGTCAATTTTCTGCAGCTAACAT 298
Db 1748 CACTCTTTTTCAGCCTGTTGGCAAGGTTCATAGGGTCAATTTTCTGCAGCTAACAT 1689
QY 299 TGAAGCCGAATAGAGGTATCGTTGGTTAACTGTTGCATAGCTCCCTGCAATTTGGTTCGG 358
Db 1688 TGAAGCCGAATAGAGGTATCGTTGGTTAACTGTTGCATAGCTCCCTGCAATTTGGTTCGG 1629
QY 359 CTGAGATTCGCCACTGNTCAAAATTTGCGACAGATTCATGTTAGGCCCTGCCACGTTGT 418
Db 1628 CTGAGATTCGCCACTGTTCAAAATTTGCGACAGATTCATGTTAGGCCCTGCCACGTTGT 1569
QY 419 TGTCTGTTGTTATGATCCACATAATAACTCTCTACGATCATCACTCTTCTTTCCCA 478
Db 1568 TGTCTGTTGTTATGATCCACATAATAACTCTCTACGATCATCACTCTTCTTTCCCA 1509
QY 479 ACTCGAGGTAAAGGTTGTTGTTCTCTCCCATGTTGGTAGTTCGAGCATTTATGACCACATA 538
Db 1508 ACTCGAGGTAAAGGTTGTTGTTCTCTCCCATGTTGGTAGTTCGAGCATTTATGATCCACATA 1449
QY 539 ATACGCTCTACCATGAGGATCTTTTCTTTGTTTCCCAACCCCTGATGCCAANGNTTCTGTGT 598
Db 1448 ATAGGTTCTACCATGAGGATCTTTTCTTTGTTTCCCA-CCCTGATGGCAAGGTTTCTGTGT 1390
QY 599 TGGCATTTCCCACTGCTGCGGTACA 624
Db 1389 TGGCATTTCCCACTGCTGCGGTACA 1364
RESULT 14
AAS80664/c
ID AAS80664 standard; cDNA; 2217 BP.
XX AC AAS80664;
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #16468.
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX Homo sapiens.
XX WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US008631.
XX
PR 31-MAR-2000; 2000US-00540217.
PR 23-AUG-2000; 2000US-00649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
DR

DR P-PSDB; ABG16477.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX
 PS Claim 1; SEQ ID NO 16468; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pt_sequences
 XX
 SQ Sequence 2217 BP; 708 A; 434 C; 505 G; 570 T; 0 U; 0 Other;
 Query Match 39.5%; Score 246.6; DB 5; Length 2217;
 Best Local Similarity 64.3%; Pred. No. 5.2e-67;
 Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
 QY 12 TAAGCATTGTGGACCCCTTTAGTTCAGATGACATCCATGCGAGGATCTTTGAATGTT 71
 DB 1079 TAGGCTATCTGAGTCCATGTTCTAGGCGAGATTTCTCTGCGGGATCTATATAGTG 1020
 QY 72 GTTGTCTTCTGTATGATCAACAAAGTACCTTACACCTTCACAGTATATCTAATTTCC 131
 DB 1019 GTAGTCTTCTATGTGGTCCCAAAATATGGAATTCATCCACCTGTGAATCTCATTTCC 960
 QY 132 CAGCCTTCTGGCAGGGGTTCTTCTCATCTGTGAAGCTTGAGTCTTCCACTGTC 191
 DB 959 CAACCTTCAGTAAGGGCTTTTCTATTAATGACCTTGACTTCTGGGGTCTTCCACTGT 900
 QY 192 GTTGTCTTCTGTATGATCAACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTTC 251
 DB 899 GTAAATCGTGTGTGTGTGACGAAATATACCTGCAATGCTGTCTGTCTCTCTCTCC 840
 QY 252 CAGCCTGCTGCAAGGTCCATAGGCTCATTTCT-----GCAGCTAAC 296
 DB 839 CATCCAGGTGCAATGACCAAGAGGATCAATTTTACTTTGTGTGATGATGAGCAATAA 780
 QY 297 ATTGAACCCCAATAGAGGTATCGTTGTGTAACTGTTGTCATAGCTCCCTGCAATGTTTC 356
 DB 779 TCTGTATCCCAATGAATCTCTGTTTAACTGCTGCAATCTCTCTGAGCTGACTA 720
 QY 357 CGCTGAGATTCCTGATCCAAATATTCGAGACAGATCCATGTTGAGCCGCTGCGCAGTT 416
 DB 719 CGCTGTAGTGCCTATGTTTCATAGTTCGAGGAGATCCAGTGTGCGCTCTGCGCAGTT 660
 QY 417 GTTGTCTTCTGTATGATCAACAAATTAACCTCTCTACGATCATCACTCTCTTCTCC 476
 DB 659 GTTGTCTTCTGTAATGTTGCAACATAAATATACGTCCTCCATGTTGTGCAACCGCGTTC 600
 QY 477 CAACCTCGAGGTAAGGTTGTGGTCTCTCCCATGTGTGTGAGCATTTATGANCACA 536
 DB 599 CAGCCAGGAGTAGAGGTTCTGGTCTATCCCAATGTTCTTTCTCAACATGATCTACA 540
 QY 537 TAATAGCTCTACATGAGGATCTTTTCTTGTTCACCCCTGATGCGCAANGTTCTGT 596
 DB 539 TAGTAAATCGCCGCTGCTGCTCCACTCTCTGCTCCCAACAGGTGCGAAGGAGCTTGA 480
 QY 597 GTT 599
 DB 479 GTT 477
 RESULT 15
 ID ABA82679/c
 XX ABA82679 standard; DNA; 2377 BP.
 AC ABA82679;
 DT 25-JAN-2002 (first entry)
 DE Atrophin-1 interacting protein (AIP4) gene SEQ ID NO:65.
 KW Human; high bone mass; HBM gene; Zmax1 gene; chromosome 11; 11q13.3;
 KW sequence tagged site; STS; osteoporosis; osteopathic; gene therapy;
 KW antisenese therapy; vaccine; bone disorder; Paget's disease; sclerostosis;
 KW osteomalacia; fibrous dysplasia; ds.
 XX Homo sapiens.
 OS WO200177327-A1.
 PN 18-OCT-2001.
 XX 21-JUN-2000; 2000WO-US016951.
 PR 05-APR-2000; 2000US-00543771.
 PR 05-APR-2000; 2000US-00544398.
 XX (GENO-) GENOME THERAPEUTICS CORP.
 XX Carulli JP, Little RD, Recker RR, Johnson ML;
 XX WPI; 2001-657171/75.
 XX New high bone mass (HBM) and Zmax1 genes and proteins useful for
 PT modulating bone mass for the treatment of e.g. osteoporosis.
 XX Claim 79; Page 367; 443pp; English.
 XX The present invention describes the human Zmax1 gene and the high bone
 CC mass (HBM) gene, which are found on chromosome 11q13.3. The Zmax1 and HBM
 CC genes have osteopathic activities. The genes can be used in gene therapy,
 CC antisenese therapy and in the production of vaccines. They can be used in
 CC the diagnosis and treatment of bone disorders including osteoporosis,
 CC Paget's disease, sclerostosis, osteomalacia and fibrous dysplasia.
 CC ABA82038 to ABA82700 and AAG68168 to AAG68193 represent sequences used in
 CC the exemplification of the present invention
 XX
 SQ Sequence 2377 BP; 765 A; 458 C; 532 G; 622 T; 0 U; 0 Other;
 Query Match 39.5%; Score 246.6; DB 5; Length 2377;
 Best Local Similarity 64.3%; Pred. No. 5.4e-67;
 Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
 QY 12 TAAGCATTGTGGACCCCTTTAGTTCAGATGATCCATGCGAGGATCTTTGAATGTT 71
 DB 1079 TAGGCTATCTGAGTCCATGTTCTAGGCGAGATTTCTCTGCGGGATCTATATAGTG 1020
 QY 72 GTTGTCTTCTGTATGATCAACAAAGTACCTTACACCTTCACAGTATATCTAATTTCC 131
 DB 1019 GTAGTCTTCTATGTGGTCCCAAAATATGGAATTCATCCACCTGTGAATCTCATTTCC 960
 QY 132 CAGCCTTCTGGCAGGGGTTCTTCTCATCTGTGAAGCTTGAGTCTTCCACTGTC 191
 DB 959 CAACCTTCAGTAAGGGCTTTTCTATTAATGACCTTGACTTCTGGGGTCTTCCACTGT 900
 QY 192 GTTGTCTTCTGTATGATCAACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTTC 251
 DB 899 GTAAATCGTGTGTGTGTGACGAAATATACCTGCAATGCTGTCTGTCTCTCTCTCC 840
 QY 252 CAGCCTGCTGCAAGGTCCATAGGCTCATTTCT-----GCAGCTAAC 296
 DB 839 CATCCAGGTGCAATGACCAAGAGGATCAATTTTACTTTGTGTGATGATGAGCAATAA 780
 QY 297 ATTGAACCCCAATAGAGGTATCGTTGTGTAACTGTTGTCATAGCTCCCTGCAATGTTTC 356
 DB 779 TCTGTATCCCAATGAATCTCTGTTTAACTGCTGCAATCTCTCTGAGCTGACTA 720
 QY 357 CGCTGAGATTCCTGATCCAAATATTCGAGACAGATCCATGTTGAGCCGCTGCGCAGTT 416
 DB 719 CGCTGTAGTGCCTATGTTTCATAGTTCGAGGAGATCCAGTGTGCGCTCTGCGCAGTT 660
 QY 417 GTTGTCTTCTGTATGATCAACAAATTAACCTCTCTACGATCATCACTCTCTTCTCC 476
 DB 659 GTTGTCTTCTGTAATGTTGCAACATAAATATACGTCCTCCATGTTGTGCAACCGCGTTC 600
 QY 477 CAACCTCGAGGTAAGGTTGTGGTCTCTCCCATGTGTGTGAGCATTTATGANCACA 536
 DB 599 CAGCCAGGAGTAGAGGTTCTGGTCTATCCCAATGTTCTTTCTCAACATGATCTACA 540

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 05:25:19 ; Search time 531 Seconds
(without alignments)
7345.241 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtntaagcattt.....tcccagactgctgcgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 6067389 seqs, 3125258755 residues

Total number of hits satisfying chosen parameters: 12134778

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq.*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq.*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq.*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq.*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq.*
18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq.*
19: /cgn2_6/ptodata/1/pubpna/US10G_PUBCOMB.seq.*
20: /cgn2_6/ptodata/1/pubpna/US10H_PUBCOMB.seq.*
21: /cgn2_6/ptodata/1/pubpna/US10I_PUBCOMB.seq.*
22: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq.*
23: /cgn2_6/ptodata/1/pubpna/US11A_PUBCOMB.seq.*
24: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq.*
25: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq.*
26: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	616	98.7	624	19	US-10-618-408-3
2	566.2	90.7	2052	14	US-10-097-534-3
3	566.2	90.7	2052	14	US-10-185-050-45
4	566.2	90.7	2052	15	US-10-101-510-56
5	566.2	90.7	2052	18	US-10-363-828-3
6	566.2	90.7	2731	17	US-10-188-186-113
7	566.2	90.7	2793	19	US-10-618-408-1

Sequence 2, Appli
Sequence 65, Appli
Sequence 383, Appl
Sequence 65, Appl
Sequence 128, Appl
Sequence 644, Appl
Sequence 644, Appl
Sequence 644, Appl
Sequence 234, Appl
Sequence 40, Appl
Sequence 1706, Ap
Sequence 4, Appli
Sequence 10, Appl
Sequence 2166, Ap
Sequence 47, Appl
Sequence 2090, Ap
Sequence 6195, Ap
Sequence 6144, Ap
Sequence 100, Appl
Sequence 25, Appl
Sequence 322, App
Sequence 7, Appli
Sequence 8, Appli
Sequence 203, App
Sequence 132, App
Sequence 30, Appl
Sequence 28, Appl
Sequence 91, Appl
Sequence 29, Appl
Sequence 1, Appli
Sequence 1640, Ap
Sequence 1640, Ap
Sequence 3, Appli
Sequence 1799, Ap
Sequence 6296, Ap
Sequence 98, Appl
Sequence 5, Appli

US-10-097-534-2
US-10-374-979-65
US-10-240-425-383
US-10-182-936A-65
US-10-283-975A-128
US-10-477-238A-644
US-10-680-287A-644
US-10-477-173-644
US-09-919-030-234
US-10-489-740-40
US-10-956-157-1706
US-10-097-534-4
US-10-363-828-10
US-10-723-860-2166
US-10-185-050-47
US-10-185-050-47
US-10-062-674-2090
US-10-723-860-6195
US-10-723-860-6144
US-09-774-639-100
US-09-969-730-25
US-10-621-363-25
US-09-764-853-322
US-10-097-534-7
US-10-097-534-8
US-10-171-581-203
US-10-473-974-132
US-10-149-827A-30
US-10-149-827A-28
US-09-764-853-91
US-10-149-827A-29
US-10-058-518-1
US-10-172-118-1640
US-10-342-887-1640
US-10-313-955-3
US-10-956-157-1799
US-10-032-585-6296
US-10-785-813-98
US-10-313-955-5

ALIGNMENTS

RESULT 1
US-10-618-408-3
; Sequence 3, Application US/10618408
; Publication No. US20040191220A1
; GENERAL INFORMATION:
; APPLICANT: Paz Einat, Louis Deiss, and Ruth Maya
; TITLE OF INVENTION: WWF1 AND USES THEREOF
; FILE REFERENCE: 67722-A; 074/US1
; CURRENT APPLICATION NUMBER: US/10/618,408
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: 60/395358
; PRIOR FILING DATE: 2002-07-11
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11)..(11)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (36)..(36)
; OTHER INFORMATION: n is a, c, g, or t
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (374)..(374)
; OTHER INFORMATION: n is a, c, g, or t

; FEATURE:		; Sequence 3, Application US/10097534	
; NAME/KEY: misc feature		; Publication No. US20030049607A1	
; LOCATION: (383)..(383)		; GENERAL INFORMATION:	
; OTHER INFORMATION: n is a, c, g, or t		; APPLICANT: GREENER, TSVIKA	
; FEATURE:		; APPLICANT: MOSKOWITZ, HAIM	
; NAME/KEY: misc feature		; APPLICANT: REISS, YUVAL	
; LOCATION: (531)..(531)		; APPLICANT: ALROY, IRIS	
; OTHER INFORMATION: n is a, c, g, or t		; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL	
; FEATURE:		; TITLE OF INVENTION: NATURATION	
; NAME/KEY: misc feature		; FILE REFERENCE: PLV-001.01	
; LOCATION: (598)..(598)		; CURRENT APPLICATION NUMBER: US/10/097,534	
; OTHER INFORMATION: n is a, c, g, or t		; CURRENT FILING DATE: 2002-03-12	
; FEATURE:		; PRIOR APPLICATION NUMBER: 60/275,224	
; NAME/KEY: misc feature		; PRIOR FILING DATE: 2001-03-12	
; LOCATION: (590)..(590)		; PRIOR APPLICATION NUMBER: 60/308,958	
; OTHER INFORMATION: n is a, c, g, or t		; PRIOR FILING DATE: 2001-07-31	
; FEATURE:		; PRIOR APPLICATION NUMBER: 60/340,170	
; NAME/KEY: misc feature		; PRIOR FILING DATE: 2001-12-07	
; LOCATION: (621)..(621)		; NUMBER OF SEQ ID NOS: 71	
; OTHER INFORMATION: n is a, c, g, or t		; SOFTWARE: PatentIn Ver. 2.1	
; US-10-618-408-3		; SEQ ID NO 3	
		; LENGTH: 2052	
		; TYPE: DNA	
		; ORGANISM: Homo sapiens	
		; US-10-097-534-3	
Query Match		90.7%; Score 566.2; DB 14; Length 2052;	
Best Local Similarity		97.3%; Pred. No. 7.6e-166;	
Matches 624; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;	
Qy	1	CCGCCGCTNTAAGCATTTGTGGACCCCTTTAGTNCAGATGACTTCCCATTCGCGAGGAT	60
Db	1		60
Qy	61	CTTTGAATGTTGTTCTTGTGTTATGATCAACAAAGTACCTTACACCTTCAACGAGTAT	120
Db	61		120
Qy	121	ATCTAATTTCCAGCCTTCGCGAGGGTCTTCATCTGTGAAGCTTCTGAGTCTTGGAT	180
Db	121		180
Qy	181	CTTCCCACTGGGTGTTTGTCTGTTATGATTACAAAGTAAACCTCTGTGTAATCCA	240
Db	181		240
Qy	241	CTCTTTTTCCTCCAGCTGTGGCAAGGTCCATAAGGGTCATTTCTGCGAGCTAAACAT	300
Db	241		300
Qy	301	AAGCCGAATAGAGTATCGTTGGTTAAACCTGTTGCATAGCTCCCTGCAATTGGTTCGGCT	360
Db	301		360
Qy	361	GAGATTGCCACTGNTCAAAATTNCGGACAGATTCATGTTAGCGCGCTGCCAGTTGTTG	420
Db	361		420
Qy	421	TTCGTGTTTATGATCCACATAATAAACTCCTCTAGCATCATCACTCTCTTTCCCAAC	480
Db	421		480
Qy	481	CTGGAGTAAAGTTGGTCTCTCCCATCTGGTGTTCGAGCATTTATGANCACATAAT	540
Db	481		540
Qy	541	ACGCTCTACCATGAGGATCTTTTCTTTGTTTCCACCCCTGATGGCAANGNTTCTGTGTTG	600
Db	541		600
Qy	601	GCAATCCAGACTGCTGCCGNACA	624
Db	601		624
RESULT 2		US-10-097-534-3/c	

; Sequence 3, Application US/10097534		Query Match	
; Publication No. US20030049607A1		Best Local Similarity	
; GENERAL INFORMATION:		97.3%; Pred. No. 7.6e-166;	
; APPLICANT: GREENER, TSVIKA		Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;	
; APPLICANT: MOSKOWITZ, HAIM		QY	
; APPLICANT: REISS, YUVAL		DB	
; APPLICANT: ALROY, IRIS		QY	
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL		DB	
; TITLE OF INVENTION: NATURATION		QY	
; FILE REFERENCE: PLV-001.01		DB	
; CURRENT APPLICATION NUMBER: US/10/097,534		QY	
; CURRENT FILING DATE: 2002-03-12		DB	
; PRIOR APPLICATION NUMBER: 60/275,224		QY	
; PRIOR FILING DATE: 2001-03-12		DB	
; PRIOR APPLICATION NUMBER: 60/308,958		QY	
; PRIOR FILING DATE: 2001-07-31		DB	
; PRIOR APPLICATION NUMBER: 60/340,170		QY	
; PRIOR FILING DATE: 2001-12-07		DB	
; NUMBER OF SEQ ID NOS: 71		QY	
; SOFTWARE: PatentIn Ver. 2.1		DB	
; SEQ ID NO 3		QY	
; LENGTH: 2052		DB	
; TYPE: DNA		QY	
; ORGANISM: Homo sapiens		DB	
; US-10-097-534-3		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	
		DB	
		QY	

; Sequence 45, Application US/10185050
; Publication No. US20030077577A1
; GENERAL INFORMATION:
; APPLICANT: Pirozzi, Gregorio
; Kay, Brian K.
; Fowlkes, Dana M.
; TITLE OF INVENTION: IDENTIFICATION AND ISOLATION OF NOVEL
; POLYPEPTIDES HAVING WW DOMAINS AND METHODS OF USING SAME
; NUMBER OF SEQUENCES: 233
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10185,050
; FILING DATE: 28-Jun-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/826,516
; FILING DATE: 03-Apr-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: MISROCK, S. LESLIE
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 1101-208-999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 896-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 45:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2052 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; SEQUENCE DESCRIPTION: SEQ ID NO: 45:
US-10-185-050-45

Query Match 90.7%; Score 566.2; DB 14; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACCTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTGTTA 86
DB 1044 CCTTTAGTTACAGATGACCTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 985

QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCTGGCAGG 146
DB 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCTGGCAGG 925

QY 147 GGTTCCTTCATTCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
DB 924 GGTTCCTTCATTCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 865

QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTTCCAGCCCTTGGTGGCAA 266
DB 864 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTTCCAGCCCTTGGTGGCAA 805

QY 267 GGTCCATTAAGGTCATTTCTGCGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
DB 804 GGTCCATTAAGGTCATTTCTGCGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

QY 327 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCCACTGNTCAAATTCGG 386
DB 744 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCCACTGNTCAAATTCGG 685

QY 387 ACAGATTCCATGTTAGCGCGCTGCCACGTTGTTGTTGTTGTTGTTATGATCCACATAATA 446
DB 684 ACAGATTCCATGTTAGCGCGCTGCCACGTTGTTGTTGTTGTTGTTATGATCCACATAATA 625

QY 447 ACTCCTCTACGATCATCACTCTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 506
DB 624 ACTCCTCTACGATCATCACTCTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 565

QY 507 CATGTGTTAGTTCGAGCATTTATGANCACACATATACGCTCTACCATGAGGATCTTTTCTT 566
DB 564 CATGTGTTAGTTCGAGCATTTATGATCCACATATAGGTTCTACCATGAGGATCTTTTCTT 505

QY 567 TGTTCACACCCCTGATGGCAANGTTCTGTGTTGGCAATTCGCCAGACTGCTGCCGNACA 624
DB 504 TGTTCACCA-CCCTGATGGCAAGGTTCTGTGTTGGCAATTCGCCAGACTGCTGCCGNACA 448

RESULT 4

US-10-101-510-56/c
; Sequence 56, Application US/10101510
; Publication No. US20030148295A1
; GENERAL INFORMATION:
; APPLICANT: WAN, JACKSON
; APPLICANT: WANG, YIXIN
; TITLE OF INVENTION: EXPRESSION PROFILES AND METHODS OF USE
; FILE REFERENCE: 15117.0012
; CURRENT APPLICATION NUMBER: US/10/101,510
; CURRENT FILING DATE: 2002-03-20
; PRIOR APPLICATION NUMBER: 60/276,947
; PRIOR FILING DATE: 2001-03-20
; NUMBER OF SEQ ID NOS: 805
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 56
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-101-510-56

Query Match 90.7%; Score 566.2; DB 15; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;

QY 27 CCCTTTAGTNCAGATGACCTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 86
DB 1044 CCTTTAGTTACAGATGACCTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTGTTA 985

QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCTGGCAGG 146
DB 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTTCTGGCAGG 925

QY 147 GGTTCCTTCATTCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
DB 924 GGTTCCTTCATTCTGTAAGCCTTCAGTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 865

QY 207 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTTCCAGCCCTTGGTGGCAA 266
DB 864 TGATTCACAAAGTAAACCTCTGTTGTAATCCACTCTTTTTTCCAGCCCTTGGTGGCAA 805

QY 267 GGTCCATTAAGGTCATTTCTGCGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
DB 804 GGTCCATTAAGGTCATTTCTGCGAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 745

QY 327 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCCACTGNTCAAATTCGG 386
DB 744 AACTGTTGCATAGCTCCCTGCAATTCGTTCCGCTGAGATTCGCCACTGNTCAAATTCGG 685

QY 387 ACAGATTCCATGTTAGCGCGCTGCCACGTTGTTGTTGTTGTTGTTATGATCCACATAATA 446
DB 684 ACAGATTCCATGTTAGCGCGCTGCCACGTTGTTGTTGTTGTTGTTATGATCCACATAATA 625

QY 447 ACTCCTCTACGATCATCACTCTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 506
DB 624 ACTCCTCTACGATCATCACTCTCTTCCCAACCTGGAGGTAAGGTTGTTGTTCTCTCC 565

Db 624 ACTCTTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGGTCTCTCC 565
QY 507 CATGTGGTAGTTCGACATTATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566
Db 564 CATGTGGTAGTTCGAGTATATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505
QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTGGCAITTCAGAGACTGCTGCCGNACA 624
Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTGGCAITTCAGAGACTGCTGCCGNACA 448
RESULT 5
US-10-363-828-3/c
; Sequence 3, Application US/10363828
; Publication No. US20040076973A1
; GENERAL INFORMATION:
; APPLICANT: Isis Pharmaceuticals, Inc.
; APPLICANT: Brett P. Monia
; APPLICANT: Lex M. Cowbert
; TITLE OF INVENTION: ANTISENSE MODULATION OF UBIQUITIN PROTEIN LIGASE EXPRESSION
; FILE REFERENCE: RTSP-0164
; CURRENT APPLICATION NUMBER: US/10/363,828
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: 09/657,481
; PRIOR FILING DATE: 2000-09-07
; NUMBER OF SEQ ID NOS: 93
; SEQ ID NO 3
; LENGTH: 2052
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2052)
US-10-363-828-3
Query Match 90.7%; Score 566.2; DB 18; Length 2052;
Best Local Similarity 97.3%; Pred. No. 7.6e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTGTTCTTCTTGTTA 86
Db 1044 CCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTGTTCTTCTTGTTA 985
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 146
Db 984 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 925
QY 147 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
Db 924 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 865
QY 207 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTGGCAAA 266
Db 864 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTGGCAAA 805
QY 267 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 326
Db 804 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 745
QY 327 AACTGTTGCATAGCTCCCTGCAATTTGGTTCCTGAGATGTCACCTGNTCAAAATTCGG 386
Db 744 AACTGTTGCATAGCTCCCTGCAATTTGGTTCCTGAGATGTCACCTGNTCAAAATTCGG 685
QY 387 ACAGATTCATGTTAGGCGCTGCCAGTTGTGTTCTGTGTTATGATCCACATAATAA 446
Db 684 ACAGATTCATGTTAGGCGCTGCCAGTTGTGTTCTGTGTTATGATCCACATAATAA 625
QY 447 ACTCTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 506
Db 624 ACTCTCTACGATCATCAACTCTTCTTCCCAACCTGGAGGTAAGGTTGTGTTCTCTCC 565
QY 507 CATGTGGTAGTTCGAGGATATGANCACATAATAGCTCTACCATGAGGATCTTTTCTT 566

Db 564 CATGTGGTAGTTCGAGTATATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTT 505
QY 567 TGTTCCTCCACCCCTGATGGCAANGNTTCTGTGTGGCAITTCAGAGACTGCTGCCGNACA 624
Db 504 TGTTCCTCA-CCCTGATGGCAAGGTTCTGTGTGGCAITTCAGAGACTGCTGCCGNACA 448
RESULT 6
US-10-188-186-113/c
; Sequence 113, Application US/10188186
; Publication No. US20040029789A1
; GENERAL INFORMATION:
; APPLICANT: Anderson et al.
; TITLE OF INVENTION: NOVEL PROTEINS AND NUCLEIC ACIDS ENCODING SAME
; FILE REFERENCE: 21402-397C
; CURRENT APPLICATION NUMBER: US/10/188,186
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: 60/303046
; PRIOR FILING DATE: 2001-07-05
; PRIOR APPLICATION NUMBER: 60/360814
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/303828
; PRIOR FILING DATE: 2001-09-07
; PRIOR APPLICATION NUMBER: 60/323380
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: 60/361133
; PRIOR FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: 60/304016
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: 60/304502
; PRIOR FILING DATE: 2001-07-11
; PRIOR APPLICATION NUMBER: 60/305262
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: 60/373881
; PRIOR FILING DATE: 2002-04-19
; PRIOR APPLICATION NUMBER: 60/305673
; PRIOR FILING DATE: 2001-07-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 368
; SOFTWARE: Custom
; SEQ ID NO 113
; LENGTH: 2731
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (11)...(2705)
US-10-188-186-113
Query Match 90.7%; Score 566.2; DB 17; Length 2731;
Best Local Similarity 97.3%; Pred. No. 8.7e-166;
Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
QY 27 CCTTTAGTNCAGATGACTTCCCATTCGAGGATCTTTGAATGTGTTCTTCTTGTTA 86
Db 1620 CCTTTAGTACAGATGACTTCCCATTCGAGGATCTTTGAATGTGTTCTTCTTGTTA 1561
QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 146
Db 1560 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGGCTTCTGGCAGG 1501
QY 147 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
Db 1500 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 1441
QY 207 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTGGCAAA 266
Db 1440 TGATTCACAAAGTAAACCTGCTGTGTGAATCCACTCTTTTCCAGGCTGTGGCAAA 1381
QY 267 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 326
Db 1380 GGTTCCTTCATTCTGTAAGCTTGAGTTCTTGGATCTTCCACCTGGGTTGTTTGTGTTA 1321

QY 327 AACTGTGATAGTCTCCCAATGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 386
 Db 1320 AACTGTGATAGTCTCCCAATGGTTCGGCTGAGATGGCCACTGTTCAAAATTCGG 1261
 QY 387 ACAGATTCCATGTAGGCGCTGCGAGTGTGTTCTGCTGATATGATCCACATAATA 446
 Db 1260 ACAGATTCCATGTAGGCGCTGCGAGTGTGTTCTGCTGATATGATCCACATAATA 1201
 QY 447 ACTCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 506
 Db 1200 ACTCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 1141
 QY 507 CATGTGGTAGTTCGAGCAATATGATCCACATAATAGCTCTACATGAGGATCTTTTCT 566
 Db 1140 CATGTGGTAGTTCGAGCAATATGATCCACATAATAGCTCTACATGAGGATCTTTTCT 1081
 QY 567 TGTTCACCCCTGATGGCAAGNTTCTGTGTTGGCATTCACAGACTGCTGCCGNACA 624
 Db 1080 TGTTCACA-CCCTGATGGCAAGNTTCTGTGTTGGCATTCACAGACTGCTGCCGTACA 1024
 RESULT 7
 US-10-618-408-1/c
 ; Sequence 1, Application US/10618408
 ; Publication No. US20040191220A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Paz Einat, Louis Deiss, and Ruth Maya
 ; FILE REFERENCE: 67722-A; 074/US1
 ; CURRENT APPLICATION NUMBER: US/10/618, 408
 ; PRIOR FILING DATE: 2003-07-11
 ; PRIOR APPLICATION NUMBER: 60/395358
 ; NUMBER OF SEQ ID NOS: 3
 ; SOFTWARE: PatentIn version 3.2
 ; SEQ ID NO 1
 ; LENGTH: 2793
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(2793)
 ; US-10-618-408-1

Query Match 90.7%; Score 566.2; DB 19; Length 2793;
 Best Local Similarity 97.3%; Pred. No. 8.8e-166;
 Matches 582; Conservative 0; Mismatches 15; Indels 1; Gaps 1;
 QY 27 CCTTTAGTACAGATGACTCCCAATGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 86
 Db 1610 CCTTTAGTACAGATGACTCCCAATGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 1551
 QY 87 TGTATCAAAAGTACCTTACACCTTACAGATATATCTAATTTCCAGGCTTCTGCGAGG 146
 Db 1550 TGTATCAAAAGTACCTTACACCTTACAGATATATCTAATTTCCAGGCTTCTGCGAGG 1491
 QY 147 GGTTCCTTCATCTGTAAAGCTTCTGAGTCTTCCCACTGCGGTTGTTTGTGTTA 206
 Db 1490 GGTTCCTTCATCTGTAAAGCTTCTGAGTCTTCCCACTGCGGTTGTTTGTGTTA 1431
 QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACCTTTTCCAGGCTGCTGCGAA 266
 Db 1430 TGATTCACAAAGTAAACCCCTGCTGTGTAATCCACCTTTTCCAGGCTGCTGCGAA 1371
 QY 267 GGTTCATAGGTCATTTTCTGAGCTTACATTTGAAGCCGAATAGAGGATCGTTGTTA 326
 Db 1370 GGTTCATAGGTCATTTTCTGAGCTTACATTTGAAGCCGAATAGAGGATCGTTGTTA 1311
 QY 327 AACTGTGATAGTCTCCCAATGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 386
 Db 1310 AACTGTGATAGTCTCCCAATGGTTCGGCTGAGATGGCCACTGNTCAAAATTNCGG 1251
 QY 387 ACAGATTCCATGTAGGCGCTGCGAGTGTGTTCTGCTGATATGATCCACATAATA 446

Db 1250 ACAGATTCCATGTAGGCGCTGCGAGTGTGTTCTGCTGATATGATCCACATAATA 1191
 QY 447 ACTCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 506
 Db 1190 ACTCTCTACGATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGCTCTCC 1131
 QY 507 CATGTGGTAGTTCGAGCAATATGATCCACATAATAGCTCTACATGAGGATCTTTTCT 566
 Db 1130 CATGTGGTAGTTCGAGCAATATGATCCACATAATAGCTCTACATGAGGATCTTTTCT 1071
 QY 567 TGTTCACCCCTGATGGCAAGNTTCTGTGTTGGCATTCACAGACTGCTGCCGNACA 624
 Db 1070 TGTTCACA-CCCTGATGGCAAGNTTCTGTGTTGGCATTCACAGACTGCTGCCGTACA 1014
 RESULT 8
 US-10-097-534-2/c
 ; Sequence 2, Application US/10097534
 ; Publication No. US20030049607A1
 ; GENERAL INFORMATION:
 ; APPLICANT: GREENER, TSVIKA
 ; APPLICANT: MOSKOWITZ, HAIM
 ; APPLICANT: REISS, YUVAL
 ; APPLICANT: ALROY, IRIS
 ; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE MODULATION OF VIRAL
 ; FILE REFERENCE: PLV-001.01
 ; CURRENT APPLICATION NUMBER: US/10/097, 534
 ; PRIOR FILING DATE: 2002-03-12
 ; PRIOR APPLICATION NUMBER: 60/275, 224
 ; PRIOR FILING DATE: 2001-03-12
 ; PRIOR APPLICATION NUMBER: 60/308, 958
 ; PRIOR FILING DATE: 2001-07-31
 ; PRIOR APPLICATION NUMBER: 60/340, 170
 ; PRIOR FILING DATE: 2001-12-07
 ; NUMBER OF SEQ ID NOS: 71
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 2
 ; LENGTH: 2377
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-097-534-2

Query Match 39.5%; Score 246.6; DB 14; Length 2377;
 Best Local Similarity 64.3%; Pred. No. 6.6e-56;
 Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
 QY 12 TAAGCAATTTGTGACCCCTTTAGTACAGATGACTTCCCAATGGAGGATCTTTGAATGTT 71
 Db 1079 TAGGCTATCTGAGGTCCTTGTCTAGGGCAGATTTTCTGTGGGGGATCTATATAGGTG 1020
 QY 72 GTTGTCTTGTGTATGATCAAAAGTACCTTACACCTTACAGATATATCTAATTTCC 131
 Db 1019 GTAGTTCCTTCTATTTGTGGTCCAAATAATATGGAATTCCTGTAATCTCAITTTCC 960
 QY 132 CAGCCTTCTGGCAGGGGTTCTTCAATCTGTAAAGCTTGAAGTCTTCCCACTGG 191
 Db 959 CAACCTTCTGAGTAAGGGCTTTTCAATTAATGACCTTGAAGTCTTCCCACTGG 900
 QY 192 GTTGTGTTTGTGTATGATTCACAAAGTAAACCCCTGTGTGTAATCCCACTCTTTTTC 251
 Db 899 GTAAATTCGTGTGTGTGTGTGACGAAATATATCTGCTGCAATGCTGTCTCTCTCTCC 840
 QY 252 CAGCCTGTTGGCAAGGTCCTAAGGTCATTTTCT-----GGAGCTAAC 296
 Db 839 CATCCAGGTGGCAATGGAACGAGGATCAAAATCTTTTACTTTGTGATGAGCAATAATA 780
 QY 297 ATTGAAGCCGAATAGAGGATCGTTGGTTAAACTGTTGATAGTCTCCCTCAATTTGTT 356
 Db 779 TCTTGATTCCCATTAATGATCTCTGTTAACTGCTGCAATGCTCTCTTGAAGCTGACTA 720
 QY 357 CGCTGAGATTGCCACTGNTCAAAATTTCCGACAGATTTCCATGTTAGGCGCTGCCAGTT 416

Db 719 CCGCTAGCTGCCATGTTTCATAGTTCGCGAGGATCCAGTGTGGCTCTGCCACGTT 660
QY 417 GTTGTCTGGTGTATGATCCATAAATAAATCTCTCTACGATCATCAACTCTTCTTTCC 476
Db 659 GTTGTCTTGTGAATGGTCAACATAAATAAATACGTCCTCCATGTTGTCAACCGCGGTTCC 600
QY 477 CAACCTGGAGGTAAAGGTGTGGTCTCTCCCATGTGGTGTAGTTCGAGCATTAATGANCACA 536
Db 599 CAGCCAGGAGGTAGAGGTCTGGTCTATCCCATGTTGTCTTTCTCAACATGATCTACA 540
QY 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTCTCCACCCCTGATGGCAANGNTTCTGT 596
Db 539 TAGTAACTCGCCGCTGCTGGTCCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 480
QY 597 GTT 599
Db 479 GTT 477

RESULT 9

US-10-374-979-65/c
; Sequence 65, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; CURRENT FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 65
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-65

Query Match 39.5%; Score 246.6; DB 17; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTTGTGACCCCTTTAGTNCAGATGACTCCCATTTGGAGGATCTTTGAATGTT 71
Db 1079 TAGGCTATCTGAGGTCCATTTGCTAGGCGAGATTTTCCTGTGGGGGATCTATATAGGTG 1020
QY 72 GTTGTCTTGTGTATGATCAACAAAGTACCTTTACACCTTTACAGAGTATATCTAAATTTCC 131
Db 1019 GTAGTCTTCTATTGTGGTCCAAATAATGGAATTCATCCACTGTGAATCTCAATTTCC 960
QY 132 CAGCCTTCTGGCAGGGGTCTTCATTTCTGTAGCCCTTGAGTCTTGGATCTTCCCACTGG 191
Db 959 CAACCTTCAGGTAAGGGCTTTTCATTTAATGACCTTGACTTCGGGCTTCCCAATGT 900
QY 192 GTTGTCTTGTGTATGATTCACAAAGTAAACCTGTGTGTAATTCACCTCTTTTTCCTCC 251
Db 899 GTAATTCGTGTGTGTGGTGTAGCAATATATCTCTGCAATGCTGTCTCTCTCTCC 840
QY 252 CAGCCTTGGTGGCAAGGTCCATAGGGTCAATTTCT-----GCAGTAAAC 296
Db 839 CATCCAGGTGGCAATGCAACCAAGGATCAAAATCTTTACTTGTGTGATGTAGCAATAAA 780
QY 297 ATTGAAGCCGAATAGAGGTATCGTGTGTAACCTGTGTGCAATAGTCTCCCTGCAATGCTTC 356

Db 779 TCTTGATTTCCATAAATGAATCTCTGGTTAAACTGCTGCAATTTGCTCTTTGAAGTGAATA 720
QY 357 CGCTGAGATTCGCACTGNTCAAAATTTCGACAGATTCATGTTAGGCGCTGCGCACGTT 416
Db 719 CGCTGAGCTGCCATTTGTTATGTTCCGAGCGGATTCAGTGTGGCTCTGCCACGTT 660
QY 417 GTTGTCTTGTGGTGTATGATCCACATAAATAAATCTCTCTACGATCATCAACTCTTCTTTCC 476
Db 659 GTTGTCTTGTGAATGGTCAACATAAATAAATACGTCCTCCATGTTGTCAACCGCGGTTCC 600
QY 477 CAACCTGGAGGTAAAGGTGTGGTCTCTCCCATGTGGTGTAGTTCGAGCATTAATGANCACA 536
Db 599 CAGCCAGGAGGTAGAGGTCTGGTCTATCCCATGTTGTCTTTCTCAACATGATCTACA 540
QY 537 TAATACGCTCTACCATGAGGATCTTTTCTTTTCTCCACCCCTGATGGCAANGNTTCTGT 596
Db 539 TAGTAACTCGCCGCTGCTGGTCCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 480
QY 597 GTT 599
Db 479 GTT 477

RESULT 10

US-10-240-425-383/c
; Sequence 383, Application US/10240425
; Publication No. US20040033502A1
; GENERAL INFORMATION:
; APPLICANT: Williams, Amanda
; APPLICANT: Boland, Joseph F.
; APPLICANT: Lord, Reginald V.
; APPLICANT: Alvarez, Chris
; APPLICANT: Wetzel, Jon C.
; APPLICANT: Scherf, Uwe
; APPLICANT: Vockley, Joseph G.
; TITLE OF INVENTION: Gene Expression Profiles in Esophageal Tissue
; FILE REFERENCE: 44921-5026
; CURRENT APPLICATION NUMBER: US/10/240,425
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: PCT/US01/09847
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 60/193,446
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 1588
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 383
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20040033502A1 AF038564
US-10-240-425-383

Query Match 39.5%; Score 246.6; DB 18; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.6e-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;
QY 12 TAAGCATTTGTGACCCCTTTAGTNCAGATGACTTCCTCCATTTGGAGGATCTTTGAATGTT 71
Db 1079 TAGGCTATCTGAGGTCCATTTGCTAGGCGAGATTTTCCTGTGGGGGATCTATATAGGTG 1020
QY 72 GTTGTCTTGTGTATGATCAACAAAGTACCTTTACACCTTCACGATATATCTAAATTTCC 131
Db 1019 GTAGTCTTCTATTGTGGTCCAAATAATGGAATTCATCCACTGTGAATCTCAATTTCC 960
QY 132 CAGCCTTCTGGCAGGGGTCTTCATTTCTGTAGCCCTTGAGTCTTGGATCTTCCCACTGG 191
Db 959 CAACCTTCAGGTAAGGGCTTTTCATTTAATGACCTTGACTTCGGGCTTCCCAATGT 900
QY 192 GTTGTCTTGTGTATGATTCACAAAGTAAACCTGTGTGTAATTCACCTCTTTTTCCTCC 251
Db 899 GTAATTCGTGTGTGTGGTGTAGCAATATATCTCTGCCAATGCTGTCTGTCTCTCTCTCC 840

Qy	252	CAGCTGTGTGGAAAGGTCCATAAGGGTCATTTTCT-----CGAGCTAAC	296
Db	839	CATCAGGTGGCAATGGACCAAGAGGATCAAAATCTTTACTTTGTGATGAGCAATAAA	780
Qy	297	ATTGAAGCCGAATAGAGGATATCGTTGGTTAAACTCTGTCATAGCTCCCTGCAATGGTTC	356
Db	779	TCTTGATTTCCCATAAATGAATCTCTGGTTAAACTGCTGCAATGCTCTTTGAAGCTGACTA	720
Qy	357	CGCTGAGATTCGCCACTGNTCAAAATTTNCGGACAGATTCCTATGTTAGCCCGCTGCCACGTT	416
Db	719	CGCTGTAGTCGCCATTGTCATAGTTCCGGACGANTCCAGTGTGGCCCTCTGCCACGTT	660
Qy	417	GTTGTTCTGTGGTTATGATCCACATAATAAACTCTCTACGATCATCAACTCTTCTTTCC	476
Db	659	GTTGTTCTTGTGAAATGGTCAACATAATAAATACGTCCCATGTTGTCAACCCGCGTTC	600
Qy	477	CAACCTCGAGCTAAAGGTTGTGGTCTCTCCCATGTTGGTAGTTTCGAGCATATATGANCACA	536
Db	599	CAGCCAGAGGTAGAGGTTCTGGTCTATCCCATGTTGTTCTTTCTCAACATGATCTACA	540
Qy	537	TAAATACGCTTACCATGAGGATCTTTTCTTTTGTCCCACTCCCTGATGGCAANENNTTCTGT	596
Db	539	TAGTAAACTCGCCCGTGTGGTCCACTCTCTGCTCCCAACCAAGGTGGCAAGGAGCTTGA	480
Qy	597	GTT	599
Db	479	GTT	477

```

RESULT 11
US-10-182-936A-65/c
; Sequence 65, Application US/10182936A
; Publication No. US20040038860A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bheem
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yavorsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; CURRENT FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-65

```

	Query Match	39.5%	Score 246.6	DB 18	Length 2377
	Best Local Similarity	64.3%	Pred. No. 6.6e-66		
	Matches 388	Conservative 0	Mismatches 200	Indels 15	Gaps 1
Qy	12	TAAGCATTTGTGGACCCCTTTAGTNCAGATGACTTCCATTGGAGAGATCTTTGAATGTT	71		
Db	1079	TAGGCTATCTGAGGTCCATTGTCTAGGGCCAGATTTCTCTGTCGGGGGATCTATATAGGTG	1020		
Qy	72	GTGTGTTCTGTGTTTATCATCAACAAAGTCACTTACACCTTCACAGATATATCTAAATTTCC	131		
Db	1019	GTATGTTCTTCTATTGTGGTCACAAAAATATGGAATTCATCCACTGTGAAATCTCATTTCC	960		
Qy	132	CAGCCTTCTCGCAGGGGTTCTTCACTCTGTAAGCCTTGGATTTCTTGATCTTCCCACGTG	191		

[illegible]

```

RESULT 12
US-10-283-975A-128/c
; Sequence 128, Application US/10283975A
; Publication No. US20040110792A1
; GENERAL INFORMATION:
; APPLICANT: Ortho-Clinical Diagnostics, Inc.
; TITLE OF INVENTION: Methods For Assessing and Treating Leukemia
; FILE REFERENCE: CDS 293 PCT
; CURRENT APPLICATION NUMBER: US/10/283,975A
; CURRENT FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/340,938
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/338,997
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/340,081
; PRIOR FILING DATE: 2001-10-30
; PRIOR APPLICATION NUMBER: 60/341,012
; PRIOR FILING DATE: 2001-10-30
; NUMBER OF SEQ ID NOS: 900
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 128
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: HUMAN
US-10-283-975A-128

```

	Query Match	39.5%	Score 246.6	DB 19	Length 2377	
	Best Local Similarity	64.3%	Pred. No. 6.6e-66			
	Matches 388	Conservative 0	Mismatches 200	Indels 15	Gaps 1	
Qy	12	TAAGCATTTGTGGACCCCTTTAGTCNCAGATGACTTCCATTTGCAGAGATCTTTGGAATGTT	71			
Db	1079	TAGGCTATCTGAGGTCCATTGTCTAGGGCAGATTTTCTGTGGGGAGTCTATATAGGTG	1020			
Qy	72	GTTGTCTCTGTGTATGATCAACAAGTAGTACCTTTACACCTTTCACGAGPATATCTAAATTTCC	131			
Db	1019	GTAGTTCTTCTCTATTGTGTCTCCACAAATATGGAAATTCATTCCTCACTGTGAATCTCATTTCC	960			

	Query Match	39.5%;	Score 246.6;	DB 20;	Length 2377;
	Best Local Similarity	64.3%;	Pred. No. 6.6e-66;		
	Matches 388;	Conservative	0;	Mismatches 200;	Indels 15; Gaps 1;
Qy	12	TAAGCATTTGTGGACCCCTTTAGTNCAGATGACTTCCCATTTGCAGAGATCTTTGAAATGTT	71		
Db	1079	TAGGCTATCTCAGGTCCATTTGTCTAGGGCAGATTTTCTGTGCGGGAGATATATAGTG	1020		
Qy	72	GTTGTTCTTTGTTTATGATCAACAAAGTACCTTACACCTTTCACAGTATATCTAAATTTCC	131		
Db	1019	GTAGTTCTTCTATTGTGTGCACAAAATATGGAATTCATCCACTGTGGAATCTCATTTTC	960		
Qy	132	CAGCCTTTCTGGCAGGGTTCTTCATTTCTGTAAGCCTTGAGTTCTTGGATCTTCCCACTGG	191		
Db	959	CAACCTTTCAGGTAAAGGGCTTTTCATTTAAATGACCTTGGACTTCTGGGGTCTTCCCATTTG	900		
Qy	192	GTTCTTTTTGTGTTTATGATTCACAAAGTAAACCCCTGTCTGTTGAATCCACTCTTTTTTTTC	251		
Db	899	GTAATTCGTGTGTTGTGGTTCACGAAATATATCTGCGCATTTGCTGTCTTCTTCTTCC	840		
Qy	252	CAGCCTGGTGGCAAGGTCCATAAGGTCATTTTTCT-----GCAGCTAAC	296		
Db	839	CATCCAGTGGCAATGCACCAAGAGGATCAAAATCTTTTACTTTGTGATGTAGCAAAATAAA	780		
Qy	297	ATTGAAGCCGAAATAGAGGTATCGTTGGTTAAACTGGTTGATAGCTCCCTCGCAATTGTTTC	356		
Db	779	TCTTGTATTTCCCATAAATGAATCTCTGTTTAAACTGCTGCATTTGCTTCTTGAAGCTGACTA	720		
Qy	357	CGCTGAGATGSCCACTGNTCAAAATTCNGGACAGATTCATGTTGAGCGCTGCACGTT	416		
Db	719	CGCTGTAGTGCCTATGTTATAGTTCGGACGGATTCAGGTGTGGCCCTCTGCCAGTT	660		
Qy	417	GTTGTTCTGGTGTATGATCCACATAATAAACTCCTCTACGATCATCAACTCTTTCTTTTC	476		
Db	659	GTTGTTCTTTGGAATGGTCAACATAATAAATACGTCCCATGTTGTCAACCCCGCGTTCC	600		
Qy	477	CAACCTGGAGGTAAAGTTGTGTTCTCTCCGATGTGTAGTTTCGAGCAATTTATGANCACA	536		
Db	599	CAGCCAGGAGGTAGAGGTTCTGGTCTATCCCATGTTGTTCTTTTCTCAACATGATCTACA	540		
Qy	537	TAATACGCTCTACCATGAGGATCTTTTCTTTGTTTCCACCCCTGATGGCAANGTTCTGT	596		
Db	539	TAGTAAACTCGCCCGTCTGTTGTCACCTCTCTGCTCTCCCAACAGGTGCAAGGAGCTTGA	480		
Qy	597	GTT 599			
Db	479	GTT 477			

RESULT 15
US-10-477-173-644/c
; Sequence 644, Application US/10477173
; Publication No. US20050070699A1
; GENERAL INFORMATION:
; APPLICANT: Genome Therapeutics Corporation and
; APPLICANT: Allen, Kristina M.
; APPLICANT: Yaworsky, Paul
; APPLICANT: Morales, Arturo J.
; APPLICANT: Graham, James R.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Liu, Wei
; TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
; FILE REFERENCE: 032796-135
; CURRENT APPLICATION NUMBER: US/10/477,173
; CURRENT FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293

```

; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 1086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 644
; LENGTH: 2377
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-477-173-644

Query Match      39.5%; Score 246.6; DB 21; Length 2377;
Best Local Similarity 64.3%; Pred. No. 6.ee-66;
Matches 388; Conservative 0; Mismatches 200; Indels 15; Gaps 1;

Qy      12  TAAGCATTTGTGGACCCCTTTAGTNCAGATGACTCCCATTTGGAGAGATCTTTGAATGTT 71
Db      1079 TAGGCTATCTGAGGTCCATTTGTCAGGCGNAGATTTCTGTGCGGGAGATCTATATAGGTG 1020

Qy      72  GTTGTTCTTTGTGTTATGATCAACAAGTAGTACCTTTACACCTTCACGAGTATATCTAAATTTCC 131
Db      1019 GTATGTTCTTCTTATTTGTGGTCCACAAAATATGGAATTTCCCATCTGGAATCTCATTTCC 960

Qy      132  CAGCCTCTCGCAGGGGTTCTTCATTTCTGTAAGCCTTTGAGTTCTTCGATCTTCCCACTGG 191
Db      959  CAACCTTCAGGTAAAGGCTTTTTCAATTAATGACCTTGACTTCTGGGGTCTTCCCAATGT 900

Qy      192  GTTGTTTTTTGTGTTATGATTCACAAAGTAAACCTCTCTGTTGAAATCCCACTCTTTTTTCC 251
Db      899  GTAATTCGTGTGTTGTGTTGACGAAATATACTCTGCCAATGCTGTCGTCTCTTCTCTCC 840

Qy      252  CAGCCTGGTGGCAAGGTCATTAAGGTCATTTTCT-----GGAGCTAAC 296
Db      839  CATCCAGGTGGCAATGGACCAAGAGGATCAAAATCTTTACTTTTGATGTAGCAAAATAA 780

Qy      297  ATTGAAGCCCAATAGAGGTATCGTTGGTTAAACTGTTGATAGTCTCCCTGCAATTCGGTTC 356
Db      779  TCTTGATCCCAATAATGAATCTCTGTTTAAACTGCTGCATTCCTCTTGAAGCTGACTA 720

Qy      357  CGCTGAGATTGCCACTGNTCAAAATTNCGGACAGATTCAATGGTAGCGCGCTGCCACGTT 416
Db      719  CGCTGTAGCTGCCAATGTTTCATAGTTCCGGACGGATTTCAGTGTGGCGCTCTGCCACGTT 660

Qy      417  GTTGTTCTGTGTTATGATCCCAATAATAAACTCTCTACGATCATCAACTCTTCTTTTCC 476
Db      659  GTTGTCTTGTGAAATGGTCAACATTAATAATAGTCCCATGTTGTCAACCCGCCGTTCC 600

Qy      477  CAACCTGGAGGTTAAAGTTGTGTGTCCTCCATGTGGTAGTTTCGAGCATTATGANCACA 536
Db      599  CAGCCAGGAGGTAGAGGTTCTGTGTCATCCCATGTTGTTCTTTTCTCAACATGATCTACA 540

Qy      537  TAATACGCTCTACCATGAGGATCTTTTCTTTGTTGCCACCCCTGTAGTGGCAAGNNTTCTGT 596
Db      539  TAGTAACCTCGCCCGTGTGGTCCACTCTCTGCTCCCAACAGGTGGCAAGGAGCTTGA 480

Qy      597  GTT 599
Db      479  GTT 477

```

Search completed: June 30, 2005, 06:39:42
Job time : 634 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 30, 2005, 04:55:15 ; Search time 2866 Seconds
(without alignments)
8287.545 Million cell updates/sec

Title: US-10-618-408-3
Perfect score: 624
Sequence: 1 ccgcgcgtnttaagcattt.....tccagactgctgcgcgnaca 624

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*

1: gb_est1.*
2: gb_est2.*
3: gb_hic.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_gss1.*
9: gb_gss2.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	525.2	84.2	899	7	CF551931	AGENCOURT
C 2	508.6	81.5	3107	3	BC026829	Mus muscu
C 3	508.6	81.5	4628	3	AK033138	Mus muscu
C 4	495.6	79.4	742	7	CN533573	UI-M-HO0-
C 5	487.6	78.1	3823	3	AK032232	Mus muscu
C 6	481.8	77.2	1113	7	CK232395	ILUUMIGEN
C 7	473.4	75.9	654	7	CK981350	BB624213
C 8	453.4	72.7	655	2	BB624213	BB624213
C 9	434.8	69.7	1101	2	BE869796	BB624213
C 10	425.8	68.2	620	4	B1598894	601446622
C 11	416	66.7	830	5	B0509208	AGENCOURT
C 12	391.6	62.8	630	1	AL049041	DKFZp434P
C 13	380.6	61.0	655	6	CB297052	12B22059
C 14	376	60.3	474	1	A1750898	cn06C12.X
C 15	373.6	59.9	557	2	BF369480	RCO-GN009
C 16	367.4	58.9	527	2	BF333137	RCO-BT081
C 17	341.2	54.7	448	5	BQ304314	RCO-BT081
C 18	339.2	54.4	469	2	BF333147	RCO-BT081
C 19	333	53.4	519	2	BF333178	RCO-BT081
C 20	329.2	52.8	731	7	CK470232	AGENCOURT
C 21	322.8	51.7	394	5	BQ304311	RCO-BT081
C 22	317.2	50.8	810	7	CK474855	AGENCOURT
C 23	309	49.5	762	6	CD802832	UI-M-GV0-
C 24	307.8	49.3	401	2	BF333140	RCO-BT081

25	307.6	49.3	470	2	BF333183	BF333183 RCO-BT081
26	300.2	48.1	445	5	BQ365503	RCO-GN009
C 27	292.4	46.9	899	7	CF583889	AGENCOURT
C 28	287.6	46.1	569	4	BG969441	602837146
C 29	286.8	46.0	687	4	B1459570	603200582
C 30	283.4	45.4	814	2	BF119329	601757344
C 31	276.6	44.3	512	4	BF173321	ic86h11.Y
C 32	271.6	43.5	563	4	B1884127	fm99h09.Y
C 33	260.4	41.7	712	6	CB457326	714668 MA
C 34	260	41.7	854	7	CV117296	AGENCOURT
C 35	251	40.2	692	7	CF747411	UI-M-HO0-
C 36	249.2	39.9	716	7	CO432363	UI-M-HO0-
C 37	249.2	39.9	1774	3	AK037940	Mus muscu
C 38	249.2	39.9	2481	3	AK048303	Mus muscu
C 39	246.6	39.5	1050	5	BM917937	AGENCOURT
C 40	244.2	39.1	704	7	CN397454	170004706
C 41	240.8	38.6	528	1	AJ727321	AJ727321
C 42	239	38.3	592	5	BP221259	BP221259
C 43	238.2	38.2	674	1	AJ727320	AJ727320
C 44	232.2	37.2	700	7	CK635728	UI-M-HO0-
C 45	231.6	37.1	581	5	BQ379860	MR4-UT009

ALIGNMENTS

RESULT 1
CF551931/c
LOCUS
DEFINITION
AGENCOURT 15595822 NIH_MGC_183 Homo sapiens cDNA clone
IMAGE:30530049 5', mRNA sequence.

ACCESSION
CF551931

VERSION
CF551931.1

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

REFERENCE
1 (bases 1 to 899)

TITLE
NIH-MGC http://mgc.nci.nih.gov/.

AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL
Unpublished (1999)

COMMENT
Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: dsgerb@xmail.nih.gov

Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)

Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov

Plate: NDAM618 row: j column: 10
High quality sequence start: 90
High quality sequence stop: 700.

FEATURES
Location/Qualifiers

1..899

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30530049"

/lab_host="DH10B-Tona (T1 and T5 phage resistant)"

/clone_lib="NIH_MGC_183"

/note="Organ: Pooled muscle (cardiac and skeletal);

Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2:

NotI; Library is oligo-dr primed and directionally cloned

(EcoRV site is destroyed upon cloning). Average insert

size 1.7. Library was constructed by Invitrogen."

ORIGIN

Query Match

84.2%; Score 525.2; DB 7; Length 899;

```

Best Local Similarity 96.8%; Pred. No. 5e-144;
Matches 544; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 27 CCCTTTAGTCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db 577 CCTTTAGTTACAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 518

QY 87 TGATCAACAAGTACTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 146
Db 517 TGATCAACAAGTACTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 458

QY 147 GGTCTTTCATCTGTAAGCCTTGAGTCTTTCGATCTTCCCATTCGAGTCTGTTGTTGTTA 206
Db 457 GGTCTTTCATCTGTAAGCCTTGAGTCTTTCGATCTTCCCATTCGAGTCTGTTGTTGTTA 398

QY 207 TGATTCACAAGTAAACCTCTGTTGAATCCACTCTTTTCCAGCCCTTCGCGAGG 266
Db 397 TGATTCACAAGTAAACCTCTGTTGAATCCACTCTTTTCCAGCCCTTCGCGAGG 338

QY 267 GGTCCATAAGGGTCAATTTCTGCACTAAACATTCGAAGCCGATAGAGTATCGTTGTTA 326
Db 337 GGTCCATAAGGGTCAATTTCTGCACTAAACATTCGAAGCCGATAGAGTATCGTTGTTA 278

QY 327 AACTGTTGCATAGCTCCCTGCAATGTTTCGCTGAGATTCGCACTGNTCAAAATTCGG 386
Db 277 AACTGTTGCATAGCTCCCTGCAATGTTTCGCTGAGATTCGCACTGNTCAAAATTCGG 218

QY 387 ACAGATTCATAGGTAGCGCGTGCACGTTGTTCTGCTGTTGTTATGATCCACATATAA 446
Db 217 ACAGATTCATAGGTAGCGCGTGCACGTTGTTCTGCTGTTGTTATGATCCACATATAA 158

QY 447 ACTCCTTCACATCATCAACTCTTCTTCCCAACTCGAGTAAAGTGTGGTCTCTCC 506
Db 157 ACTCCTTCACATCATCAACTCTTCTTCCCAACTCGAGTAAAGTGTGGTCTCTCC 98

QY 507 CATGTG-GTAGTTCGAGATATGACATATATGATCGTCTTACCATGAGATCTTTTCT 565
Db 97 CATGTGNTAGTTCGAGTATATGATCCACATATAGTCTTACCATGAGATCTTTTCT 38

QY 566 TTGTTTCCACCCCTTGATGGCAA 587
Db 37 TTGTTCCACCCCTTGATGGCAA 16

RESULT 2
BC026829/c
LOCUS
DEFINITION
Mus musculus WW domain containing E3 ubiquitin protein ligase 1,
mRNA (CDNA clone IMAGE:4209397), containing frame-shift errors.
ACCESSION
BC026829.1 GI:20071294
VERSION
HTC.
KEYWORDS
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 3107)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udwin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loguercio, N.A., Peters, G.J.,
Abramson, K.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.M., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,

```

```

Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalhus, D.E.,
Schnerch, A., Schein, J.E., Jones, S.O. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 3107)
Strausberg, R.
Direct Submission
Submitted (04-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps-i@mail.nih.gov
Tissue Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loussegh, H.,
Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 37 Row: h Column: 15
This clone has the following problem: frame shifted.
FEATURES
Location/Qualifiers
1..3107
/organism="Mus musculus"
/mol_type="mRNA"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:4209397"
/tissue_type="Colon, normal. 5 month old male mouse."
/clone_lib="NCI CGAP Co24"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
ORIGIN
Query Match 81.5%; Score 508.6; DB 3; Length 3107;
Best Local Similarity 91.3%; Pred. No. 5.6e-139;
Matches 546; Conservative 0; Mismatches 51; Indels 1; Gaps 1;
QY 27 CCCTTTAGTCAGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 86
Db 1657 CCTTTAGTTACTGATGACTTCCCATTCGAGGATCTTTGAATGTTGTTGTTCTTGTGTTA 1598

QY 87 TGATCAACAAGTACTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 146
Db 1597 TGATCAACAAGTACTTACACCTTCACGAGTATATCTAAATTCACGACCTTCGCGAGG 1538

QY 147 GGTCTTTCATCTGTAAGCCTTGAGTCTTTCGATCTTCCCATTCGAGTCTGTTGTTGTTA 206
Db 1537 GGTCTTTCATCTGTAAGCCTTGAGTCTTTCGATCTTCCCATTCGAGTCTGTTGTTGTTA 1478

QY 207 TGATTCACAAGTAAACCTCTGTTGAATCCACTCTTTTCCAGCCCTTCGCGAGG 266
Db 1477 TGATTCACAAGTAAACCTCTGTTGAATCCACTCTTTTCCAGCCCTTCGCGAGG 1418

QY 267 GGTCCATAAGGGTCAATTTCTGCACTAAACATTCGAAGCCGATAGAGTATCGTTGTTA 326
Db 1417 GGTCCATAAGGGTCAATTTCTGCACTAAACATTCGAAGCCGATAGAGTATCGTTGTTG 1358

QY 327 AACTGTTGCATAGCTCCCTGCAATGTTTCGCTGAGATTCGCACTGNTCAAAATTCGG 386
Db 1357 AACTGTTGCATAGCTCCCTGCAATGTTTCGCTGAGATTCGCACTGNTCAAAATTCGA 1298

```


207 TGATTCACAAAGTAACCCCTGCTGTGTTGAATCCACTCTTTTCCAGCCTGGTGGCAAA 266
 515 TGATTCACAAAGTAACCTGCTGTGTTGAATCCACTCTTTTCCAGCCTGGTGGTAAG 456
 267 GGTCCATAAGCGGTCATTTCTGAGCTAACTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
 455 GGTCCATAAGCGGTCATTTCTGAGCTAACTGGAAGCCGAATAGAGGTATCGTTGGTTG 396
 327 AACTGTTGATAGCTCCTCCCAATTTGGTTCGGTGGATTCGACATGCTGNTCAAAATTCGG 386
 395 AATGTTGATAGCTCCTCCCAATTTGAATTCGGTGGATTCGACATGCTGNTCAAAATTCGA 336
 387 ACAGATTCATAGTGGAGCCGCTGCCAGCTGTTGTTCTGTTGATGATCAATCAATAAA 446
 335 ACAGATTCATAGTGGAGCCGCTGCCAGCTGTTGTTCTGTTGATGATCAATCAATAAA 276
 447 ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTGTGTTGTTCTCTCC 506
 275 ACTCCTCTAGCATCATCACTCTTTTCCCACTGGAGGTAAAGTGTGTTGTTCTCTCC 216
 507 CATGTGTAGTTCGAGCATATGATCCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 215 CATGTAGTATTCAGATTTATGATCCACATATAGTTCATCATGAGGATCTTTTCTC 156
 567 TGTTCACACCCCTGATGGCAAGNTTCTGTGTTGGCATTCGCCAGACTGCTGCCGNACA 624
 155 TGTTCACA-CCCTGATGGCAAGCTCTGTGTTGGTGTTCAGACTGTGGCCGCGAGA 99

RESULT 4
 .CN533573/c
 LOCUS
 DEFINITION
 UT-M-HO0-CPY-D-14-0-UI.r1 NIH BMAP_H00 Mus musculus cDNA clone
 IMAGE:30658453 5', mRNA sequence.
 CN533573
 EST.
 CN533573.1 GI:46861729
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 742)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs@email.nih.gov
 Tissue Procurement: Dr. James Lin University of Iowa
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: Distribution information can be found at
 http://genome.uiowa.edu/distribution/mousefl.html
 This clone was contributed by the Brain Molecular Anatomy Project
 (BMAP)

Seq primer: pYX-5.
 Location/Qualifiers
 1. 742
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAG:30658453"
 /tissue_type="Upper Head"
 /dev_stage="9.5-10.5 dpc"
 /lab_host="DH10B (T1 phage
 /clone_lib="NIH BMAP H00"
 /notes="Organ: Head; Vector: pYX-Asc; Site 1: Eco I;
 Site 2: Not I; The library was constructed according
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose
 gel. First strand cDNA synthesis was primed with oligo-dT
 primer containing a Not I site. Double strand cDNA was

size selected according to mRNA size fraction, ligated
 with EcoR I adaptor, digested with NotI and then cloned
 directionally into pYX-Asc vector. The library tag
 sequence located between the Not I site and the polyA tail
 is CGACCTGAAT. This library was created for the University
 Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the
 Developing Mouse Nervous System', supported by National
 Institute of Mental Health (NIMH)."

ORIGIN

Query Match 79.4%; Score 495.6; DB 7; Length 742;
 Best Local Similarity 90.8%; Pred. No. 2.7e-135;
 Matches 544; Conservative 0; Mismatches 53; Indels 2; Gaps 2;

QY 27 CCCTTTAGTNCAGATGACATCTCCCATTTGCAGAGATCTTTGAATGTTCTGTTCTTGTGTTA 86
 639 CCTTTAGTACTGATGACATCTCCCATTTGCAGAGGTCTNTGAATGTTGTTCTGCTGTTAN 580
 QY 87 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTCGCAGG 146
 579 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGCCTTCTCGCAGG 520
 QY 147 GGTCTCTCATCTGTAAGCCTTGAGTCTTGAGATCTTCCACTGGGTTGTTTGTGTTA 206
 519 GGTCTCTCATCTGTAAGCCTTGAGTCTTGAGATCTTCCACTGGGTTGTTTGTGTTA 460
 QY 207 TGATTCACAAAGTAAACCCCTGCTGTGTTGAATCCACTCTTTTCCAGCCTGCTGGCAAA 266
 459 TGATTCACAAAGTAAACCTGCTGTGTTGAATCCACTCTTTTCCAGCCTGCTGGTGAAG 400
 QY 267 GGTCCATAAGGGTCAATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTA 326
 399 GGTCCATAAGGGTCAATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGTTG 340
 QY 327 AACTGTTGATAGCTCCCTGCGAATTTGGTTCGCTGAGATTCACATGTCATAAATTCGG 386
 339 AATGTTGATAGCTCCCTGCGAATTTGAATTCGCTGAGACTGCCACTGTTCAAAATTCGA 280
 QY 387 ACAGATTCATAGTGGCGCTGCCAGCTGTTGTTCTGCTGTTATATGATCCACATAATA 446
 279 ACAGATTCATAGTGGCGCTGCCAGCTGTTGTTCTGCTGTTATATGATCCACATAATA 220
 QY 447 ACTCCTCTACGATCATCACTCTTTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCC 506
 219 ACTCCTCCAGTCACTCACTCTTTTCCCACTGGAGGTAAAGGTTGTGTTCTCTCC 160
 QY 507 CATGTTGATTCGAGCATTTATGACCACATATAGCTCTACCATGAGGATCTTTTCTT 566
 159 CATGTTGATTCGAGCATTTATGATCCACATAATAGGTTCTACCATGAGGATCTTTTCTC 100
 QY 567 TGTTCACACCCCTGATGGCAAGNTTCTGTTGG-CAITTCACAGACTGCTGCCGNACA 624
 99 TGTTCACA-CCCTGATGGCAAGCTCTGTTGGTGTTCAGACTGTGGCCGCGAGA 42

RESULT 5
 AK032232/c
 LOCUS
 DEFINITION
 Mus musculus adult male olfactory brain cDNA, RIKEN full-length
 enriched library, clone:6430508G12 product:WMP1 (FRAGMENT) homolog
 [Homo sapiens], full insert sequence.
 AK032232
 HTC; CAP trapper.
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1
 Carninci P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 99279253


```

10349636
2
REFERENCE
AUTHORS
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
TITLE
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
JOURNAL
MEDLINE
20499374
PUBMED
11042159
REFERENCE
3
AUTHORS
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuami,T., Taehiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsuura,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and Hayashizaki,Y.
TITLE
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
JOURNAL
MEDLINE
20530913
PUBMED
11076861
REFERENCE
4
AUTHORS
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
TITLE
Functional annotation of a full-length mouse cDNA collection
JOURNAL
MEDLINE
Nature 409, 685-690 (2001)
PUBMED
11076861
REFERENCE
5
AUTHORS
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
TITLE
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
JOURNAL
MEDLINE
Nature 420, 563-573 (2002)
PUBMED
12038100
REFERENCE
6
AUTHORS
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Oheato,N.,
Okazaki,Y., Saito,R., Satoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,P., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
TITLE
Direct Submission
JOURNAL
MEDLINE
Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216)
COMMENT
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/.
FEATURES
source
Location/Qualifiers
1. .3823
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="FANTOM DB:6430508G12"
/db_xref="taxon:10090"
/clone="6430508G12"
/sex="male"
/tissue_type="olfactory brain"
/clone_lib="RIKEN full-length enriched mouse cDNA library"

```

```

/dev_stage="adult"
misc_feature
275..2706
/notes="WNP1 (FRAGMENT) homolog [Homo sapiens]
(SPTR|O00307, evidence: FASTV, 95.6%ID, 83.1%length,
match=1698)
putative"
polyA_signal
3797..3802
/notes="putative"
polyA_site
3823
/notes="putative"
ORIGIN
Query Match 78.1%; Score 487.6; DB 3; Length 3823;
Best Local Similarity 92.2%; Pred. No. 9.7e-133;
Matches 511; Conservative 0; Mismatches 43; Indels 0; Gaps 0;
QY 27 CCCTTTAGTNCAGATGACATCTCCCATTTGCCAGAGGATCTTTGAATGTTGTTGTTGTTGTTA 86
Db 1547 CCTTTAGTACTGATGATCTTCCCATTTGCCAGAGGCTTTGNAATGTTGTTGTTGTTGTTA 1488
QY 87 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGAGCCTTCTGGCAGG 146
Db 1487 TGATCAACAAAGTACCTTACACCTTCCAGAGTATATCTAATTTCCAGAGCCTTCTGGCAGG 1428
QY 147 GGTTCCTTCATCTGTAAGCCTTGAGTCTTGGATCTTCCACCTGGTGTGTTGTTGTTA 206
Db 1427 GGTTCCTTCATCTGTAAGCCTTGAGTCTTGGATCTTCCACCTGGTGTGTTGTTGTTA 1368
QY 207 TGATTCACAAAGTAAACCCCTGCTGTTGAATCCACTCTTTTCCAGAGCCTGGTGGCAA 266
Db 1367 TGATTCACAAAGTAAACTGCTGTTGTAATCCACTCTTTTCCAGAGCCTGGTGGTAAG 1308
QY 267 GGTTCATAAGGGTCATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTA 326
Db 1307 GGTTCATAGGGTCATTTTCTGCAGCTAACATTTGAAGCCGAATAGAGGTATCGTTGGTTG 1248
QY 327 AACTGTTGATAGTCTCCCTGCAATTTGGTTCGGTGGATGATCCACTGNTCAAAATTCGG 386
Db 1247 AATTGTTGATAGTCTCCCTGCAATTTGATTTCCGCTGAGACTGCCACTGTTTCAAAATTCGA 1188
QY 387 ACAGATTCCATGTAGCGCTGCTCCACGCTGTTGTTGTTCTGTTGTTATGATGCACATAATA 446
Db 1187 ACAGATTCCATGTAGCGCTGCTCCACGCTGTTGTTGTTCTGTTGTTATGATGCACATAATA 1128
QY 447 ACTCTCTACGATCATCACTCTTTTCCCAACCTGGAGGTAAAGGTTGTGTTCTCTCC 506
Db 1127 ACTCTCCAGATCATCACTCTTTTCCCAACCTGGAGGCAAGGCTGTGTTCTCTCC 1068
QY 507 CATGTGGTGTTCGAGCATTTAGNCACATATAGCTCTTACCATGAGGATCTTTTCTT 566
Db 1067 CATGTAGTATTCGAGTATTATGATCCACATATAGGTTTCTACCATGAGGATCTTTTCTC 1008
QY 567 TGTTCACACCCCTG 580
Db 1007 TGTTCACACCCCTG 994
RESULT 6
CK232395/c
LOCUS
CK232395.1 1113 bp mRNA linear EST 09-DEC-2003
DEFINITION
IULUMIGEN_MQ_4036 Katze_MMP12 Macaca mulatta cDNA 5' similar to
human WNP1 (Hs.191668), mRNA sequence.
ACCESSION
CK232395
VERSION
CK232395.1 GI:39638753
KEYWORDS
EST.
SOURCE
Macaca mulatta (rhesus monkey)
ORGANISM
Macaca mulatta
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
Cercopitheciinae; Macaca.
1 (bases 1 to 1113)
REFERENCE
1. (bases 1 to 1113)
AUTHORS
Katze,M.G., Thomas,M., Korth,M., Iadonato,S.P. and Magness,C.L.
TITLE
Large-scale Rhesus Macaque cDNA sequencing

```

JOURNAL
COMMENT

Unpublished (2003)
Contact: C. Magness
Illumigen Biosciences Inc.
2203 Airport Way S, Suite 450, Seattle, WA 98134, USA
Tel: 2063780400
Fax: 2063780408
Email: cmagness@illumigen.com
Sequenced on 2003.12.02. 760 Q20 bases. Assemblies in contig w/ 1 member(s). Contig contains 1 (0%) lib members.
PCR Primers
FORWARD: CCCTCACTAAGGGAACAAA
BACKWARD: CACTATAGGCGAATGGGTA
Insert Length: 1113 Std Error: 0.00
Plate: CL000052 row: C column: 04
Seq primer: CCCTCACTAAGGGAACAAA
POLYA=Yes.

Location/Qualifiers
1. .1113
/organism="Macaca mulatta"
/mol_type="mRNA"
/strain="Indian"
/db_xref="taxon:9544"
/sex="male"
/dev_stage="newborn infant"
/lab_host="E. coli SOLR"
/clone_lib="Katze MMPL2"
/note="Organ: placenta; Vector: Uni-ZAP XR; Site 1: Ecor I; Site 2: Xho I; Created from Stratagene ZAP-CDNA Synthesis kit (catalog #200400) and ZAP-CDNA Gigapack III Gold Cloning Kit (Catalog #200450)."

FEATURES
source

ORIGIN

Query Match 77.2%; Score 481.8; DB 7; Length 1113;
Best Local Similarity 95.5%; Pred. No. 3.7e-131;
Matches 515; Conservative 0; Mismatches 21; Indels 3; Gaps 2;

QY 27 CCCTTTAGTCAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 86
DB 539 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 480

QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 146
DB 479 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 420

QY 147 GGTTCCTCATCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
DB 419 GGTTCTTCATCTGTAAAGCCTTGAGTCTTTGGATCTTCCACCTGGGTTGTTTGTGTTG 360

QY 207 TGATTCACAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCCC--AGCCCTGGTGGA 264
DB 359 TGATTCACAAGTAAACCCCTGCTGTGTAATCCACTCTTTTCTCCCGCAGCGCTGGTGGA 300

QY 265 AAGTCCATAGGTCATTTCTTCGACCTAACATTTGAAGCCGAATAGAGTATCGTTGT 324
DB 299 AAGTCCATAGGTCATTTCTTCGACCTAACATTTGAAGCCGAATAGAGTATCGTTGT 240

QY 325 TAACTGTTGCATAGTCCCTGCAATGGTTCCGCTGAGATTGCCACTGNATCAAAATTTC 384
DB 239 TAACTGTTGCATAGTCCCTGCAATGGTTCCGCTGAGATTGCCACTGTTCAAAATTTC 180

QY 385 GGACAGATTCATGTTAGGCGCTGCCAGTGTCTTCTGGTGTATGATCCACATAT 444
DB 179 GGACAGATTCATGTTAGGCGCTGCCAGTGTCTTCTGGTGTATGATCCACATAT 120

QY 445 AAACTCCTCPACGATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGTCCT 504
DB 119 AAACTCCTCAGCATCATCAACTCTTTTCCCAACCTGGAGGTAAGGTTGTGTCCT 60

QY 505 -CCCATGTGTAGTTCGAGCATTTATGANCACATATATAGCTCTACCATGAGGATCTTT 562
DB 59 CCCCATGTGTAGTTCGAGCATTTATGATCCACATATATAGTTCATACCATGAGGATCTTT 1

RESULT 7

CK981350/c
LOCUS
DEFINITION
CK981350 654 bp mRNA linear EST 16-MAR-2004
4113416 BARC 9BOV Bos taurus cDNA clone 9BOV42_007 5', mRNA
sequence.
CK981350
ACCESSION
VERSION
CK981350.1 GI:45499330
KEYWORDS
EST.
SOURCE
Bos taurus (cow)
ORGANISM
Bos taurus
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE
1 (bases 1 to 654)
Sonstegard, T.S., Van Tassel, C.P., Matukumalli, L.K., Harhay,
G.P., Bosak, S., Rubenfield, M. and Gasbarre, L.C.
Production of EST from cDNA libraries derived from immunologically
activated bovine gut
Unpublished (2004)
CONTACT: Tad S. Sonstegard
Bovine Functional Genomics Laboratory
Animal and Natural Resources Institute
Bd19. 200 Rm2A BARC-East, Beltsville, MD 20705, USA
Tel: 3015048416
Fax: 3015048414
Email: tads@anri.barc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
0.000925 using options -trim alt -trim fasta. Vector identified
by cross match using options -minmatch 12 -minscore 12
Plate: 42 row: O column: 07
Seq primer: CCCAGTCACGACGTTGTAAACG
High quality sequence stop: 654.

FEATURES
source

1. 654
/organism="Bos taurus"
/mol_type="mRNA"
/strain="Holstein"
/db_xref="taxon:9913"
/clone="9BOV42_007"
/sex="Male"
/tissue_type="Pooled"
/dev_stage="Multiple"
/lab_host="DH10B T1 phage resistant"
/clone_lib="BARC 9BOV"
/note="Organ: Abomasum; Vector: pagen-1; Site 1: EcorV;
Site 2: NotI; Equimolar amounts of mRNA extracted from
fundic and pyloric abomasums of 18 and 21 week old steers.
Exposure to Osteragia osteragi was initiated at 15 weeks
of age. fundic and pyloric abomasum"

ORIGIN

Query Match 75.9%; Score 473.4; DB 7; Length 654;
Best Local Similarity 87.6%; Pred. No. 9.9e-129;
Matches 524; Conservative 0; Mismatches 73; Indels 1; Gaps 1;

QY 27 CCCTTTAGTCAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 86
DB 629 CCCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTGTGTTA 570

QY 87 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 146
DB 569 TGATCAACAAAGTACCTTACACCTTCACGAGTATATCTAATTTCCAGCCCTCTGGCAGG 510

QY 147 GGTTCCTCATCTGTAAAGCCTTGAGTCTTCCACTCTCCAGTGTGTTTGTGTTA 206
DB 509 GGTTCCTCATCTGCAAACTTGAGTCTTGGGTCCTCCACTGGTGTGTTTGTGTTA 450

QY 207 TGATTCACAAGTAAACCCCTGCTGTGTAATCCACTCTTTTTCACGCTGGTGGA 266
DB 449 TGATTCACAAGTAAACCTGCTGTGTAATCCACTCTTTTTCACGCTGGTGGA 390

QY 267 GGTTCATAGGTCATTTTCTGCAGCTAACATTCGAAGCCGAATAGAGTATCGTTGTTA 326

Db 389 GGCCGGTATGGGTCATTTCTGGCGCTAAACATTTGAAGCCGAGTAGAGGTATCGCTGGTTG 330

Qy 327 AACTGTTGTCATAGCTCCCTGCAATTTGGTTCCGCTGAGATTGCCATGTCNTCAAAATTNCGG 386

Db 329 AACTGCTGCATAGCTCCCTGCAGCTGGCTCCGCTGGGACTGCCACTGCTCAAAAGTTCCGC 270

Qy 387 ACAGATTCATAGTAGGCGCCCTGCCAGCTGGTTGTTCTGTTGTTATGATGCCATATATA 446

Db 269 ACCGACTCCATAGTAGGCGCCCTGCCAGCTGGTTGTTCTGTTGTTATGTTGTTGTTAC 210

Qy 447 ACTCTCTACGATCATCACTCTTTCTTCCCACTGGAGTAAAGTTGTTGTTCTCTCC 506

Db 209 ACTTGGCCCGGTCATGACTCTTTCTTCCCACTGGAGGAGAGCTGGCCCTCTCC 150

Qy 507 CATGTGTAGTTTCGAGCATTTATGANCACATAATACGCTCTACCATGAGGATCTTTCTT 566

Db 149 CAGTGTGTGCTGGTGTATGATCCACGTAGTAGTTCTACCATGAGGATCTTTCTC 90

Qy 567 TGTTCACCCCTGATGGCAANGTTCTGTGTTGGCATTTCCAGACTGCTGCGGNACA 624

Db 89 TGTTCACA-CCCTGATGGCAGCGTTCTGTGTTGGCACTCCAGACTGCTGCGGACA 33

RESULT 8

BB624213/c

LOCUS

DEFINITION

musculus cDNA clone 8030445B08 5', mRNA sequence.

ACCESSION

BB624213

VERSION

BB624213.1

KEYWORDS

GI:15398222

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

REFERENCE

1. (bases 1 to 655)

Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toyota, T., Muramatsu, M. and Hayashizaki, Y. RIKEN Mouse ESTs [Arakawa, T., et al. 2001]

TITLE

Unpublished (2001)

JOURNAL

Contact: Yoshihide Hayashizaki

COMMENT

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Teurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-resgsc.riken.jp, URL: http://genome.gsc.riken.jp/

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.* 10 (10), 1617-1630 (2000)

wagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)

Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)

Yamanaka, I., Kiyosawa, H., Kondo, S., Saito, T., Shinagawa, A., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K., Arakawa, T., Ishii, Y. and Hayashizaki, Y.

Mapping of 19032 mouse cDNAs on mouse chromosomes. *J. Struct. Func. Genomics* 2 pre, L72-L86 (2001)

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

e mouse tissues.

FEATURES

Location/Qualifiers

1..655

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="8030445B08"

/sex="male"

/tissue_type="testis"

/dev_stage="15 days embryo"

/lab_host="DH10B"

/clone_lib="RIKEN full-length enriched, 15 days embryo male testis"

/notes="Site_1: Sali; Site_2: BamHI; cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. 1st strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGACTCTTTTCTTTTCTTTTCTTTT 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 185.0 Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGATTCGAGTTAATTAATTAATCCCTCCCCCCCC 3']. cDNA was cloned into the XhoI and BamHI sites. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: Sali; 3' end: BamHI"

ORIGIN

Query Match 72.7%; Score 453.4; DB 2; Length 655;

Best Local Similarity 90.3%; Pred. No. 8.3e-123;

Matches 503; Conservative 0; Mismatches 52; Indels 2; Gaps 2;

Qy 69 GTTGTGTTCTTGTGTATGATCAACAAAGTACCTTACACCTT-CAGGATATATTAAT 127

Db 654 GTGGTGGTCTCGGGGTTAAGTACCANAAGTACCTTCCCGAGTATACCTGAT 595

Qy 128 TTCCAGCCTTCTGGCAGGGGTTCTTCTTCTTAAGCCTTGAGTCTTCTGATCTTCCCA 187

Db 594 TTCCAGCCTTCTGGCAGGGGTTCTTCTTCTTAAGCCTTGAGTCTTCTGATCTTCCCA 535

Qy 188 CTGGGTGTTGTTTGTGTTATGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTT 247

Db 534 CTGGGTGTTGTTTGTGTTATGATTCACAAAGTAAACCTGTCTGTGAATCCACTCTTTT 475

Qy 248 TTCCAGCCTTCTGGCAGAAAGTCCATAAGGTCAATTTCTGAGCTTAACATTAAGCCGA 307

Db 474 TTCCAGCCTTCTGGTGGTGAAGGTCATATGAGGGGTCAATTTCTGCAAGTAAACATGAAGCCGA 415

Qy 308 ATAGAGGTATCGTTGTTAAACCTGTTGATAGCTCCCTGCAATTTGGTTCCGCTGAGATTG 367

Db 414 ATAGAGGTATCGTTGTTGAAATTTGTCATAGCTCCCTGCAATTTGATTTCCGCTGAGACTG 355

Qy 368 CCACGTNTCAAAATTTNCGGACAGATTCCCATGGTAGGCGCTGCCAGCTGTGTTCTCGGT 427

Db 354 CCACGTNTCAAAATTTNCGGACAGATTCCCATGGTAGGCGCTGCCAGCTGTGTTCTCGGT 295

Qy 428 GTTATGATCCACATATAAAGTCTCTACATCATCACTCTTCTTCCCAAGCTGGAGG 487

Db 294 GTTATGATCCACATATAAAGTCTCTCCAGATCATCAACTCTTCTTCCCAAGCTGGAGG 235

Qy 488 TAAAGGTGTGGTCTCTCCCATGTGTGTAGTTCCGAGCATTTATGANCACATATAACGCTCT 547

Db 234 CAAAGGCTGTGGTCTCTCCCATGTGTGTAGTTCCGAGCATTTATGATCCACATATAAGTCTCT 175

```

Qy 548 ACCATGAGGATCTTTCTTTTGTCTCCACCCCTCATGCGCAANGNTTCTGTGTTGGCATTC 607
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
174 ACCATGAGGATCTTTCTTTTGTCTCCACCCCTCATGCGCAANGNTTCTGTGTTGGTTC 116
Qy 608 CAGACTGCTGCCGNACA 624
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
115 CAGACTGTGGCCGCAGA 99

RESULT 9
BE869796/c
LOCUS BE869796
DEFINITION 1101 bp mRNA linear EST 20-OCT-2000
ACCESSION BE869796
VERSION BE869796
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9570 row: k column: 05
High quality sequence stop: 614.
Location/Qualifiers
1. 1101
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3850876"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
FEATURES
source
1. 1101
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3850876"
/tissue_type="adenocarcinoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_65"
/notes="Organ: colon; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.8 kb. Library constructed by Life
Technologies."
ORIGIN
Query Match 69.7%; Score 434.8; DB 2; Length 1101;
Best Local Similarity 95.7%; Pred. No. 3e-117;
Matches 486; Conservative 0; Mismatches 18; Indels 4; Gaps 4;
Qy 120 TATCTAATTTCCAGCCTTCTGGC-AGGGGTCTTCTCATCTGTAAAGCCTTGAGTCTTGG 178
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
664 TCTAATATGTCGAGCCTTCTGCCAGGGTCTTCTCATCTGTAAAGCCTTGAGTCTTGG 605
Qy 179 ATCTTCCAC-TGGTGTGTTTTGTGTATGATCA-AGTAAACCTGTCTGTGAA 236
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
604 ATCTTCCACGTGGTGTGTTTTGTGTATGATTCACAGAAGTAAACCTGTCTGTGAA 545
Qy 237 TCCACTCTTTTTCAGCCTGTGGCAAGGTCCATAGGGTCAATTTCTGCACCTAAC 296
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
544 TCCACTCTTTTTCAGCCTGTGGCAAGGTCCATAGGGTCAATTTCTGCACCTAAC 485
Qy 297 ATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTGTGATAGCTCCCTGCAATTTGGTTC 356
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
484 ATTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTGTGATAGCTCCCTGCAATTTGGTTC 425
Qy 357 CGCTGAGATTGCCACGTGNTCAAAATTCGACAGATTCATGTTAGGGCGCTGCCAGCTT 416
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||

```

```

Db 424 CGCTGAGATTGCCACTGTTCAAAATTTGGACAGATTTCCATGTCGCGCGCTGCCAGTT 365
Qy 417 GTTGTCTCTGTTGTATGATCCACATAATAAATCTCTCTAGATCATCAACTCTTTCTTCC 476
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
364 GTTGTCTCTGTTGTATGATCCACATAATAAATCTCTCTAGATCATCAACTCTTTCTTCC 305
Qy 477 CAACCTGGAGGTAAAGGTTGTGGTCTCTCCCATGTGGTAGTTCGAGCATTAATGATCCACA 536
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
304 CAACCTGGAGGTAAAGGTTGTGGTCTCTCCCATGTGGTAGTTCGAGCATTAATGATCCACA 245
Qy 537 TAATAGCTCTACCATGAGATCTTTCTTTGTTCCTCCACCCCTGATGGCAANGNTTCTGT 596
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
244 TAATAGCTCTACCATGAGATCTTTCTTTGTTCCTCCACCCCTGATGGCAANGNTTCTGT 186
Qy 597 GTTGGCATTTCCAGACTGCTGCCGNACA 624
Db ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
185 GTTGGCATTTCCAGACTGCTGCCGTACA 158

```

```

RESULT 10
BI598894/c
LOCUS BI598894
DEFINITION 620 bp mRNA linear EST 07-SEP-2001
ACCESSION BI598894
VERSION BI598894
KEYWORDS mRNA sequence.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11759 row: a column: 01
High quality sequence start: 7
High quality sequence stop: 575.
Location/Qualifiers
1. 620
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5299848"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/clone_lib="NIH_MGC_96"
/notes="Organ: brain; Vector: pBluescriptR (modified
pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI
(gcgag); Oligo-dt primed using primer
5'-TTTTTTTTTTTTTTVN-3', size-selected for average
insert size 2.3 kb and normalized to 10^5. This is a
primary library enriched for full-length clones and
constructed using the Cap-trapper method (Carninci, in
preparation). Library constructed by M. Brownstein
(NIH/NHGRI, National Institutes of Health). Note: this is
a NIH_MGC Library."

```

FEATURES

source

1. 620

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5299848"

/tissue_type="hypothalamus"

/lab_host="DH10B"

/clone_lib="NIH_MGC_96"

/notes="Organ: brain; Vector: pBluescriptR (modified

pBluescript KS+); Site 1: BamHI; Site 2: SalI-XhoI

(gcgag); Oligo-dt primed using primer

5'-TTTTTTTTTTTTTTVN-3', size-selected for average

insert size 2.3 kb and normalized to 10^5. This is a

primary library enriched for full-length clones and

constructed using the Cap-trapper method (Carninci, in

preparation). Library constructed by M. Brownstein

(NIH/NHGRI, National Institutes of Health). Note: this is

a NIH_MGC Library."

ORIGIN

Query Match

Best Local

Similarity

97.6%;

Pred. No. 1.2e-114;

Matches 441;

Conservative

0;

Mismatches 10;

Indels 1;

Gaps 1;

```
QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTTGTTGTTA 86
Db 460 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTTGTTGTTA 401
QY 87 TGATCAACAAGTACCTTTACACCTTTACAGAGTATATCTAATT-TCCAGAGCTTCTGCGCAG 145
Db 400 TGATCAACTAAGTACCTTTACACCTTTACAGAGTATATCTAATTGTTCCAGGCTTCTGCGCAG 341
QY 146 GGGTCTTCTTCTGTAAGCCTTTGAGTCTTTGGATCTTTCCACACTGGGTGTTGTTTGTGTT 205
Db 340 GGGTCTTCTTCTGTAAGCCTTTGAGTCTTTGGATCTTTCCACACTGGGTGTTGTTTGTGTT 281
QY 206 ATGATTCACAAGTAACCCCTGCTGTTGATTCACACTCTTTTTCAGGCTGTTGTCGCA 265
Db 280 ATGATTCACAAGTAACCCCTGCTGTTGATTCACACTCTTTTTCAGGCTGTTGTCGCA 221
QY 266 AGGTCCATAAGGGTCAATTTCTGCGAGCTAACTTGAAGCCGATAGAGGTATCGTTGGTT 325
Db 220 AGGTCCATAAGGGTCAATTTCTGCGAGCTAACTTGAAGCCGATAGAGGTATCGTTGGTT 161
QY 326 AAATCTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATTCGCCACTGNTCAAAATTCG 385
Db 160 AAATCTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATTCGCCACTGNTCAAAATTCG 101
QY 386 GACAGATTCATGTTAGGCGCTGCGACAGTTGTTGTTCTGTTGTTATGATCCACATATA 445
Db 100 GACAGATTCATGTTAGGCGCTGCGACAGTTGTTGTTCTGTTGTTATGATCCACATATA 41
QY 446 AACTCTCTACGATCATCACTCTTTCTTCCC 477
Db 40 AACTCTCTACGATCATCACTCTTTCTTCCC 9
```

```
RESULT 11
BU509208/c
LOCUS BU509208 891 bp mRNA linear EST 12-SEP-2002
DEFINITION AGENCOURT_10099965 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:6503763
5', mRNA sequence.
ACCESSION BU509208
VERSION BU509208.1 GI:22815441
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 891)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM14061 row: p column: 04
High quality sequence stop: 649.
Location/Qualifiers
1. .891
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6503763"
/tissue_type="leiomyosarcoma"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_71"
/note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 2.1 kb."
```

FEATURES
source

ORIGIN

```
Query Match 66.7%; Score 416; DB 5; Length 891;
Best Local Similarity 98.1%; Pred. No. 1.le-111;
Matches 419; Conservative 0; Mismatches 8; Indels 0; Gaps 0;
QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTTGTTGTTA 86
Db 427 CCTTTAGTTACAGATGACTTCCCATTCGCGAGGATCTTTGAATGTTGTTGTTCTTTGTTGTTA 369
QY 87 TGATCAACAAGTACCTTTACACCTTTACAGAGTATATCTAATT-TCCAGAGCTTCTGCGCAG 146
Db 367 TGATCAACAAGTACCTTTACACCTTTACAGAGTATATCTAATT-TCCAGAGCTTCTGCGCAG 308
QY 147 GGTCTTCTTCTGTAAGCCTTTGAGTCTTTGGATCTTTCCACACTGGGTGTTGTTTGTGTTA 206
Db 307 GGTCTTCTTCTGTAAGCCTTTGAGTCTTTGGATCTTTCCACACTGGGTGTTTGTGTTA 248
QY 207 TGATTCACAAGTAAACCCCTGCTGTTGATTCACACTCTTTTTCAGGCTGTTGTCGCA 266
Db 247 TGATTCACAAGTAAACCCCTGCTGTTGATTCACACTCTTTTTCAGGCTGTTGTCGCA 188
QY 267 GGTCCATAAGGGTCAATTTCTGCGAGCTAACTTGAAGCCGATAGAGGTATCGTTGTTGTTA 326
Db 187 GGTCCATAAGGGTCAATTTCTGCGAGCTAACTTGAAGCCGATAGAGGTATCGTTGTTGTTA 128
QY 327 AACTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATTCGCCACTGNTCAAAATTCG 386
Db 127 AACTGTTGCATAGCTCCCTGCAATGGTTCGCTGAGATTCGCCACTGNTCAAAATTCG 68
QY 387 ACAGATTCATGTTAGGCGCTGCGACAGTTGTTGTTCTGTTGTTATGATCCACATATA 446
Db 67 ACAGATTCATGTTAGGCGCTGCGACAGTTGTTGTTCTGTTGTTATGATCCACATATA 8
QY 447 ACTCTCTC 453
Db 7 ACTCTCTC 1
```

```
RESULT 12
AL049041/c
LOCUS AL049041 630 bp mRNA linear EST 04-SEP-2003
DEFINITION DKFP434P0218_r1.434 (synonym: htes3) Homo sapiens cDNA clone
DKFP434P0218, mRNA sequence.
ACCESSION AL049041
VERSION AL049041.1 GI:4728350
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 630)
AUTHORS Ottenwaelder, B., Obermaier, B., Mewes, H.W., Gassenhuber, J. and
Wiemann, S.
TITLE EST (Ottenwaelder, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: MIPS
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by MediGenomix (Martinsried/Germany) within the cDNA
sequencing consortium of the German Genome Project. No s1 sequence
available.
This clone (DKFP434P0218) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1. .630
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFP434P0218"
```

```
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site 1: NotI; Site 2: SalI"

ORIGIN
Query Match 62.8%; Score 391.6; DB 1; Length 630;
Best Local Similarity 98.3%; Pred. No. 1.6e-104;
Matches 394; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTTCTTGTGTTA 86
Db 401 CCTTTAGTTACAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTTCTTGTGTTA 342
QY 87 TGATCAACAAGTACTTACACCTTCACGAGTATATCTTAATTTCCAGCCTTCGCGAGG 146
Db 341 TGATCAACAAGTACTTACACCTTCACGAGTATATCTTAATTTCCAGCCTTCGCGAGG 282
QY 147 GGTCTTCTCATCTGTGAAGCCTTGAGTCTTTCGATCTTCCCACTGGGTTGTTTGTGTTA 206
Db 281 GGTCTTCTCATCTGTGAAGCCTTGAGTCTTTCGATCTTCCCACTGGGTTGTTTGTGTTA 222
QY 207 TGATCAACAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGCCTGTGGGCAAA 266
Db 221 TGATCAACAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGCCTGTGGGCAAA 162
QY 267 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCGGAATAGAGTATCGTTGGTTA 326
Db 161 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCGGAATAGAGTATCGTTGGTTA 102
QY 327 AACTGTTGCTAGTCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTNCGG 386
Db 101 AACTGTTGCTAGTCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTNCGG 42
QY 387 ACAGATTCATAGTGGAGCGCTGCGACGTTGTTGTTCTGCT 427
Db 41 ACAGATTCATAGTGGAGCGCTGCGACGTTGTTGTTCTGCT 1

RESULT 13
CB297052/c
LOCUS
DEFINITION 12B22059_rev_1_C01_r_002.ab1 Chimpanzee brain library Koo's Pan troglodytes cDNA clone 12B22059_rev_1_C01_r_002.ab1 5', mRNA sequence.
CB297052
CB297052.1 GI:28622482
EST.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE
1 (bases 1 to 655)
Hellmann,I., Zollner,S., Enard,W., Ebersberger,I., Nickel,B. and Paabo,S.
Selection on human genes as revealed by comparisons to chimpanzee cDNA
Genome Res. (2003) In press
Contact: Paabo S
Evolutionary Genetics
Max-Planck-Institute for evolutionary Anthropology
Deutscher Platz 6, 04103 Leipzig, Germany
Tel: +49- (0)-341-3550 500
Fax: +49- (0)-341-3550 555
Email: paabo@eva.mpg.de
Seq primer: M13 reverse.
Location/Qualifiers
1. 655
/organism="Pan troglodytes"
/mol_type="mRNA"
/db_xref="taxon:9598"
/clone="12B22059_rev_1_C01_r_002.ab1"

FEATURES
source
```

```
/sex="male"
/tissue_type="brain, presumably cortex"
/dev_stage="adult"
/lab_host="Epicurian Coli (TM) XL-10-Gold"
/clone_lib="Chimpanzee brain library Koo's"
/note="Vector: pUCHi; Site 1: SfiI-A; Site 2: SfiI-B; The library was prepared using the SMART cDNA library construction kit (Clontech), doing only primer extension, but not PCR amplification of the cDNA. The only deviation from the published protocol was that we cloned the cDNA into a plasmid vector."

ORIGIN
Query Match 61.0%; Score 380.6; DB 6; Length 655;
Best Local Similarity 98.2%; Pred. No. 3e-101;
Matches 383; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 27 CCCTTTAGTNCAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTTCTTGTGTTA 86
Db 393 CCTTTAGTTACAGATGACTTCCCATTCGAGAGATCTTTGAATGTTGTTTCTTGTGTTA 334
QY 87 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTTAATTTCCAGCCTTCGCGAGG 146
Db 333 TGATCAACAAGTACCTTACACCTTCACGAGTATATCTTAATTTCCAGCCTTCGCGAGG 274
QY 147 GGTCTTCTCATCTGTGAAGCCTTGAGTCTTTCGATCTTCCCACTGGGTTGTTTGTGTTA 206
Db 273 GGTCTTCTCATCTGTGAAGCCTTGAGTCTTTCGATCTTCCCACTGGGTTGTTTGTGTTA 214
QY 207 TGATCAACAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGCCTGTGGGCAAA 266
Db 213 TGATCAACAAGTAAACCTGTCTGTGAATCCACTCTTTTCCAGCCTGTGGGCAAA 154
QY 267 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCGGAATAGAGTATCGTTGGTTA 326
Db 153 GGTCCATAAGGGTCATTTCTGACGCTAAACATTGAAGCGGAATAGAGTATCGTTGGTTA 94
QY 327 AACTGTTGCTAGTCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTNCGG 386
Db 93 AACTGTTGCTAGTCTCCCTGCAATTTGGTTCGGCTGAGATTGCCACTGNTCAAAATTNCGG 34
QY 387 ACAGATTCATAGTGGAGCGCTGCGACGTT 416
Db 33 ACAGATTCATAGTGGAGCGCTGCGACGTT 4

RESULT 14
AI750898/c
LOCUS
DEFINITION AI750898 474 bp mRNA linear EST 20-JUN-2002
clone NHTBC_cn06c12 random, mRNA sequence.
AI750898
AI750898.1 GI:5129162
EST.
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 474)
Jia,L.B., Young,M.F., Touchman,J.W., Bouffard,G.G., Beckstrom-Sternberg,S.M., Green,E.D., Powell,J.I., Yang,L.M., Robey,P.G., Hotchkiss,R.N. and Francomano,C.A.
SGAP: The Skeletal Genome Anatomy Project
Unpublished (1997)
Contact: Libin Jia
Medical Genetics Branch
National Human Genome Research Institute
10/10C101, 9000 Rockville Pike, Bethesda, MD 20892-1267, USA
Tel: 301-402-4877
Fax: 301-496-7157
Email: libin@helix.nih.gov
DNA Sequencing and analyses by National Institutes of Health Intramural Sequencing Center (NISC).
```

Plate: 06 row: c column: 12
Seq primer: -21M13 forward primer (ABI).
Location/Qualifiers
1. .474
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NHTBC cn06c12"
/sex="Female"
/tissue_type="Bone"
/cell_type="Trabecular Bone Cells"
/lab_host="SURE"
/clone_lib="Normal Human Trabecular Bone Cells"
/notes="Organ: Hip; Vector: pBluescript; Site: 1; EcoRI; Library constructed by Dr. Marian Young and Dr. Pamela Gehron Robey (NIDCR)"

ORIGIN
Query Match 60.3%; Score 376; DB 1; Length 474;
Best Local Similarity 97.0%; Pred. No. 6.3e-100;
Matches 390; Conservative 0; Mismatches 11; Indels 1; Gaps 1;
QY 223 CCCTGCTGTGTAATCCACCTCTTTTCCAGAGCTGTGGCAAGGTCCATAAGGGTTCAT 282
Db 474 CCTGTCTGTGTAATCCACCTCTTTTCCAGAGCTGTGGCAAGGTCCATAAGGGTTCAT 415
QY 283 TTCTGCGAGTAACTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTC 342
Db 414 TTCTGCGAGTAACTGAAGCCGAATAGAGGTATCGTTGGTTAAACTGTTGCATAGCTC 355
QY 343 CTGCAATTTGGTCCGCTGAGATGCCACTGNTCAAAATTCGACAGATTCATCGGTAG 402
Db 354 CTGCAATTTGGTCCGCTGAGATGCCACTGNTCAAAATTCGACAGATTCATCGGTAG 295
QY 403 GCCGCTGCCAGTGTGTCTGTGTATGATCCACATATAAATTCCTCTAGATCAT 462
Db 294 GCCGCTGCCAGTGTGTCTGTGTATGATCCACATATAAATTCCTCTAGATCAT 235
QY 463 CAACCTCTCTTTCCCACTGGAGTAAAGGTGTGTCTCTCCCATGTGTAGTTCGAG 522
Db 234 CAACCTCTCTTTCCCACTGGAGTAAAGGTGTGTCTCTCCCATGTGTAGTTCGAG 175
QY 523 CATATGANCACATAATACGCTTACCATGAGATCTTTCTTTTGTTCACACCCCTGAT 582
Db 174 TATTATGATCCACATAATAGTGTCTACCATGAGATCTTTCTTTTGTTCACA-CCCTGAT 116
QY 583 GGCAANGTTCGTGTGGCATTCACAGACTGTGCCGNACA 624
Db 115 GGCAAGTTCGTGTGGCATTCACAGACTGTGCCGTACA 74

RESULT 15
BF369480
LOCUS
DEFINITION
RC0-GN0090-260900-032-e09 GN0090 Homo sapiens cDNA, mRNA sequence.
ACCESSION
BF369480
VERSION
BF369480.1 GI:11331505
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 557)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663

10737800
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=RC0&t2=RC0-GN0090-
260900-032-e09&t3=2000-09-28&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 16
High quality sequence stop: 557.
Location/Qualifiers
1..557
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="GN0090"
/note="Organ: Placenta normal; Vector: puc18; Site: 1;
SmaI; Site: 2; SmaI; A mini-library was made by cloning
products derived from ORESTES PCR (U.S. Letters Patent
application No. 196,716 - Ludwig Institute for Cancer
Research) profiles into the pUC 18 vector. Reverse
transcription of tissue mRNA and cDNA amplification were
performed under low stringency conditions."

ORIGIN
Query Match 59.9%; Score 373.6; DB 2; Length 557;
Best Local Similarity 97.7%; Pred. No. 3.4e-99;
Matches 388; Conservative 0; Mismatches 7; Indels 2; Gaps 1;
QY 27 CCCTTAGTNCAGATCACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 86
Db 161 CCTTAATACAGATCACTTCCCATTCGCGAGGATCTTGAATGTTGTTCTTGTTA 220
QY 87 TGATCAACAAAGTACCTTACACCTTCACAGTATATCTAATTTCCACAGCTTCTGGCAGG 146
Db 221 TGATCAACAAAGTACCTTACACCTTCGCGAGTATATCTAATTTCCACAGCTTCTGGCAGG 280
QY 147 GGTTCCTTCATCTGTAAAGCCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 206
Db 281 GGTTCCTTCATCTGTAAAGCCTTGAGTTCCTGGATCTTCCACCTGGGTTGTTTGTGTTA 340
QY 207 TGATCAACAAAGTAAACCTGTCTGTGAATCCACTC--TTTTTCCACGCTGTGGCA 264
Db 341 TGATCAACAAAGTAAACCTGTCTGTGAATCCACTCTTTTTTCCACGCTGTGGCA 400
QY 265 AAGTCCATAAGGTCATTTTCTGCAGCTAACCAATGAAGCCGAATAGAGGTATCGTTGGT 324
Db 401 AAGTCCATAAGGTCATTTTCTGCAGCTAACCAATGAAGCCGAATAGAGGTATCGTTGGT 460
QY 325 TAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGNTCAAAATTC 384
Db 461 TAAACTGTTGCATAGCTCCCTGCAATTTGGTTCGCTGAGATTCGCCACTGTTCAAAATTC 520
QY 385 GGACAGATTCATGGTAGCCGCTGCCACGTTGTTGT 421
Db 521 GGACAGATTCATGGTAGCCGCTGCCACGTTGTTGT 557

Search completed: June 30, 2005, 06:29:09
Job time : 2875 secs

THIS PAGE BLANK (USPTO)